ANALYSIS I & II & III

VERSION: December 5, 2022

ARMIN SCHIKORRA

Contents

In Analysis
there are no theorems
only proofs

These lecture notes take great inspiration from the lecture notes by Michael Struwe (Analysis III, German), as well as by Piotr Hajłasz (Analysis I). We will also follow the presentations in Evans-Gariepy [Evans and Gariepy, 2015] (measure theory), Grafakos [Grafakos, 2014] (Fourier Analysis) and wikipedia. Further sources are Piotr Hajlasz' Functional Analysis, Clasons [Clason, 2020] and everything available on the internet. Sometimes we follow those sources verbatim.

Pictures that were not taken from sources mentioned above (or wikipedia) are usually made with geogebra.

References

[Adams and Fournier, 2003] Adams, R. A. and Fournier, J. J. F. (2003). *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition.

[Bandle, 2017] Bandle, C. (2017). Dido's problem and its impact on modern mathematics. *Notices Amer. Math. Soc.*, 64(9):980–984.

[Bethuel, 1991] Bethuel, F. (1991). The approximation problem for Sobolev maps between two manifolds. *Acta Math.*, 167(3-4):153–206.

[Bethuel, 2020] Bethuel, F. (2020). A counterexample to the weak density of smooth maps between manifolds in Sobolev spaces. *Invent. Math.*, 219(2):507–651.

[Brezis, 2011] Brezis, H. (2011). *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York.

[Cazenave and Haraux, 1998] Cazenave, T. and Haraux, A. (1998). *An introduction to semilinear evolution equations*, volume 13 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York. Translated from the 1990 French original by Yvan Martel and revised by the authors.

[Clason, 2020] Clason, C. ([2020] ©2020). *Introduction to functional analysis*. Compact Textbooks in Mathematics. Birkhäuser/Springer, Cham.

[Evans, 2010] Evans, L. C. (2010). *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[Evans and Gariepy, 2015] Evans, L. C. and Gariepy, R. F. (2015). *Measure theory and fine properties of functions*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition.

[Gagliardo, 1957] Gagliardo, E. (1957). Caratterizzazioni delle tracce sulla frontiera relative ad alcune classi di funzioni in n variabili. *Rend. Sem. Mat. Univ. Padova*, 27:284–305.

[Gilbarg and Trudinger, 2001] Gilbarg, D. and Trudinger, N. S. (2001). *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin. Reprint of the 1998 edition.

[Grafakos, 2014] Grafakos, L. (2014). *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics*. Springer, New York, third edition.

[Hajłasz and Liu, 2010] Hajłasz, P. and Liu, Z. (2010). A compact embedding of a Sobolev space is equivalent to an embedding into a better space. *Proc. Amer. Math. Soc.*, 138(9):3257–3266.

[Hang and Lin, 2001] Hang, F. and Lin, F. (2001). Topology of Sobolev mappings. *Math. Res. Lett.*, 8(3):321–330.

[Leoni, 2017] Leoni, G. (2017). *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[Lieb and Loss, 2001] Lieb, E. H. and Loss, M. (2001). *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition.

[Maz'ya, 2011] Maz'ya, V. (2011). *Sobolev spaces with applications to elliptic partial differential equations*, volume 342 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, augmented edition.

[Megginson, 1998] Megginson, R. E. (1998). *An introduction to Banach space theory*, volume 183 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

[Mironescu, 2005] Mironescu, P. (2005). Fine properties of functions: an introduction. Lecture.

[Mironescu, 2018] Mironescu, P. (2018). The role of the Hardy type inequalities in the theory of function spaces. *Rev. Roumaine Math. Pures Appl.*, 63(4):447–525.

[Schikorra et al., 2017] Schikorra, A., Spector, D., and Van Schaftingen, J. (2017). An L^1-type estimate for Riesz potentials. *Rev. Mat. Iberoam.*, 33(1):291–303.

[Shor, 1997] Shor, P. W. (1997). Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509.

[Simon, 1996] Simon, L. (1996). *Theorems on regularity and singularity of energy minimizing maps*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel. Based on lecture notes by Norbert Hungerbühler.

[Sokal, 2011] Sokal, A. D. (2011). A really simple elementary proof of the uniform boundedness theorem. *Amer. Math. Monthly*, 118(5):450–452.

[StackExchange, a] StackExchange. Direct approach to the closed graph theorem. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/138005 (version: 2017-04-13).

[StackExchange, b] StackExchange. Orthogonality in different inner products. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/2649128 (version: 2018-02-13).

[Stein and Shakarchi, 2003] Stein, E. M. and Shakarchi, R. (2003). *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ. An introduction.

[Talenti, 1976] Talenti, G. (1976). Best constant in Sobolev inequality. *Ann. Mat. Pura Appl. (4)*, 110:353–372.

[Ziemer, 1989] Ziemer, W. P. (1989). *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York. Sobolev spaces and functions of bounded variation.

Index

Part 1. Analysis I: Measure Theory

1. Measures, σ-Algebras

A measure is a way to measure (hence the name!) volumes. So for some set X it should be a map

$$\mu : 2^X \to [0,\infty]$$

that to a subset $A \subset X$ assigns the volume $\mu(A)$. Here 2^X denotes the *power set* of X, i.e. the collection of subsets of X.

$$2^X = \{A : \quad A \subset X\}.$$

What would we want from a volume in $\mathbb{R}^n$? Well it seems to be a reasonable assumption to axiomatically assume the following

- For any $A \subset \mathbb{R}^n$ we have $\mu(A) \in [0,\infty]$
- (Invariance under translation and rotation) For any set $A \subset \mathbb{R}^n$, any rotation $P \in O(n)$ and any vector $x \in \mathbb{R}^n$ we have $\mu(x + OA) = \mu(A)$ where we denote
$$x + OA := \{x + Oa \in \mathbb{R}^n : \quad a \in A\}$$
- For any $A, B \subset \mathbb{R}^n$ disjoint we have $\mu(A \cup B) = \mu(A) + \mu(B)$

As reasonable as that sounds, there are two problems here:

- For $n \geq 3$ the only map $\mu : 2^{\mathbb{R}^n} \to [0,\infty]$ that satisfies our axiom is constant (Hausdorff, 1914)
- For $n = 1,2$ there are indeed nonconstant maps $\mu : 2^{\mathbb{R}^n} \to [0,\infty]$ that satisfy the above axioms, however even if we fix $\mu([0,1]^n) := 1$ there is more than one possibility for such a μ (Banach 1923).
- the whole business about disjoint sets is really tricky, as illustrated by the *Banach-Tarski-Paradoxon* (1924):

  Let $n \geq 3$, A and B be bounded sets with $int(A)$ and $int(B) \neq \emptyset$. Then there exist finitely many $(x_i)_{i=1}^N \subset \mathbb{R}^n$, $(O_i)_{i=1}^N \subset O(n)$ and parwise disjoint sets $(C_i)_{i=1}^N$ so that $(x_i + O_iC_i)_{i=1}^N$ are pairwise disjoint and

$$A = \bigcup_{i=1}^N C_i, \quad \text{and} \quad B = \bigcup_{i=1}^N (x_i + O_iC_i)\,. \tag{1.1}$$

  That is we can deconstruct any set A in $\mathbb{R}^n$ into disjoint sets, move them around (without any scaling!) and obtain another completely different set B - see Figure 1.1.

This is crazy, so the axiomatic definition of a reasonable volume in $\mathbb{R}^n$ has failed, and we are back to square one.

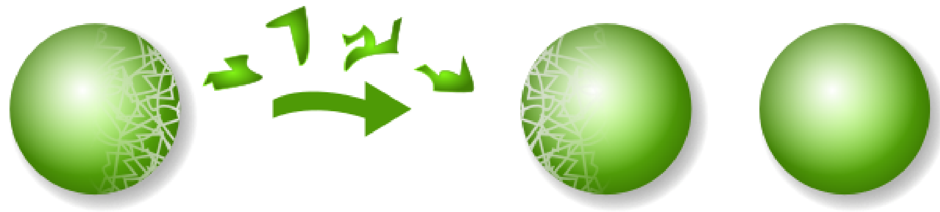


FIGURE 1.1. A ball can be decomposed into a finite number of disjoint sets and then reassembled into two balls identical to the original

So instead of defining a volume in $\mathbb{R}^n$ axiomatically, let us generally define what a reasonable notion of a volume should satisfy. Later we will then construct the Lebesgue measure that has most of the desired properties on $\mathbb{R}^n$.

Clearly $\mu(\emptyset) = 0$ is a reasonable assumption. Ideally we would also like $\mu(A \cup B) = \mu(A) \cup \mu(B)$ – but this will be a surprisingly tricky, confusing, and paradox assumption, so let us settle for the following notion

Definition 1.1. Let X be any set and 2^X the potential set of X. A map $\mu : 2^X \to [0, \infty]$ is a *measure* on X if we have

(1) $\mu(\emptyset) = 0$
(2) $\mu(A) \leq \sum_{k=1}^{\infty} \mu(A_k)$ whenever $A, A_k \subset X$, $k \in \mathbb{N}$ and $A \subset \bigcup_{k \in \mathbb{N}} A_k$

Remark 1.2. Condition (1) and (2) implies *monotonicity*,

$$\mu(A) \leq \mu(B) \quad \forall A \subset B.$$

(simply set $A_1 := B$ and $A_k := \emptyset$ for $k \geq 2$).

In particular we could equivalently replace (2) above by *σ-subadditivity*, namely

$$\mu(\bigcup_{k=1}^{\infty} A_k) \leq \sum_{k=1}^{\infty} \mu(A_k).$$

Remark 1.3.
- A word of warning: we will use here the notion of an *outer measure* that is defined on all of 2^X, not only on its σ-algebra of measurable sets.
- In particular we have $\mu(A \cup B) \leq \mu(A) + \mu(B)$ for any set $A, B \subset X$. However, in general, we cannot hope for all disjoint sets A and B hope that $\mu(A \cup B) = \mu(A) + \mu(B)$ (see above), this will lead to the notion of *non-measurable sets*.

Example 1.4 (Jordan content)**.**
- The outer *Jordan content* $J_*(E)$ of a set $E \subset \mathbb{R}^n$ is defined as follows.

  For a product of bounded cubes $C = [a_1, b_1) \times [a_2, b_2) \times \ldots \times [a_n, b_n)$ we set

$$\mathrm{vol}(C) := (b_1 - a_1) \cdot (b_2 - a_2) \cdot \ldots (b_n - a_n).$$

$$J^*(E) := \inf \left\{ \sum_{i=1}^{N} \mathrm{vol}(C_i) \quad \text{for some } N \in \mathbb{N}, \text{ and cubes } (C_i)_{i=1}^N \text{ such that } E \subset \bigcup_{i=1}^{N} C_i \right\}$$

Here we follow the convention that $\inf \emptyset = +\infty$.

$J^\varepsilon(\cdot)$ is *not* a measure: take any enumeratotion of $\mathbb{Q} \cap [0,1] = \{q_1, \ldots, q_n, \ldots\}$. Set $A_k := \{q_k\}$ and $A := \bigcup_{k=1}^\infty A_k = [0,1] \cap \mathbb{Q}$. If $(C_i)_{i=1}^N$ is a finite cover of $[0,1] \cap \mathbb{Q}$ then $\bigcup_i \overline{C_i} \supset [0,1]$[1], so $J^*(A) = 1$. However $J^*(A_k) = 0$ for each k, we have $J^*(A) \not\leq \sum_{k=1}^\infty J^*(A_k)$.

However J^* satisfies finite additivity,

$$\mu(A \cup B) \leq \mu(A) + \mu(B),$$

i.e.

$$\mu(A) \leq \sum_{k=1}^{N} \mu(A_k) \quad \text{whenever } A, A_k \subset X,\ k \in \{1, \ldots, N\},\ N \in \mathbb{N}, \text{ and } A \subset \bigcup_{k \in \mathbb{N}} A_k.$$

Such a map $J^\varepsilon : 2^X \to [0, \infty)$ is called a *content*.

- The countable version of the outer Jordan content, is called the *Lebesgue outer measure*

$$m^*(E) := \inf \left\{ \sum_{i=1}^{\infty} \operatorname{vol}(C_i) \quad \text{for some , and cubes } (C_i)_{i=1}^\infty \text{ such that } E \subset \bigcup_{i=1}^{\infty} C_i \right\} \tag{1.2}$$

It is again clear that $m^*(\emptyset) = 0$. Let now $A \subset \bigcup_{k=1}^n A_k$. We may assume that $m^*(A_k) < \infty$ otherwise there is nothing to show. Fix $\varepsilon > 0$. For each k we can pick $(C_{k;i})_{i=1}^\infty$ such that $\bigcup_{i=1}^\infty C_{k;i} \supset A_k$ and

$$\sum_{i=1}^{\infty} \operatorname{vol}(C_{k,i}) \leq m^*(A_k) + \frac{\varepsilon}{2^k}.$$

Now $\bigcup_{k,i \in \mathbb{N}} C_{k,i} \supset A$ and thus

$$m^*(A) \leq \sum_{k,i \in \mathbb{N}} \operatorname{vol}(C_{k,i}) \leq \sum_{k=1}^{\infty} m^*(A_k) + \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k}$$

That is, we have shown that for any $\varepsilon > 0$,

$$m^*(A) \leq \sum_{k=1}^{\infty} m^*(A_k) + \varepsilon$$

Taking $\varepsilon \to 0$ we conclude that $m^*(A) \leq \sum_{k=1}^\infty m^*(A_k)$ – that is $m^*(A)$ is indeed a measure.

Later the Lebesgue measure $\mathcal{L}^n$ will coincide with $m^*(A)$.

- The *inner Jordan content*,

$$J_*(E) := \sup \left\{ \sum_{i=1}^{N} \operatorname{vol}(C_i) \quad \text{for some } N \in \mathbb{N}, \text{ and cubes } (C_i)_{i=1}^N \text{ such that } \bigcup_{i=1}^{N} C_i \subset E \right\}$$

Here we follow the convention that $\sup \emptyset = 0$.

[1] Indeed, take $r \in [0,1]$ then there exists q_k converging to r, q_k belongs infinitely often to the same interval, so $r \in \overline{C_i}$ for some i

Still $J_*(\cdot)$ is not a measure. Take $A_1 := [0,1]\backslash\mathbb{Q}$ and for $i \geq 2$ we set $A_i = \{q_i\}$ for $\{q_2, \ldots, \} = \mathbb{Q} \cap [0,1]$ any enumeration of $\mathbb{Q} \cap [0,1]$. Since A_1 has empty interior we have $J_*(A_1) = 0$. Similarly, $J_*(A_i) = 0$ for $i \geq 2$. However $A := \bigcup_{i=1}^\infty A_i = [0,1]$ satisfies $J_*([0,1]) = 1$. So we have $J_*(A) \not\leq \sum_{i=1}^n J_*(A_i)$.

- If we simply make the innter Jordan content countable, i.e. if we set
$$\tilde{J}_*(E) := \sup \left\{ \sum_{i=1}^\infty \mathrm{vol}(C_i) \quad \text{for cubes } (C_i)_{i=1}^N \text{ such that } \bigcup_{i=1}^\infty C_i \subset E \right\}$$
we run into the same problem as for J_*, namely $J_*([0,1]\backslash\mathbb{Q}) = 0$. So $\tilde{J}_*(E)$ is still not a measure.

Example 1.5 (Counting measure)**.** Let X be any set. Then $\#2^X \to \mathbb{N} \cup \{0\}$ defined by
$$\#A := \text{number of elements in } A,$$
is a measure, called the *counting measure*.

Exercise 1.6. *Let X be a metric space and $\mu : 2^X \to [0,\infty]$ a measure. Let $A \subset X$ then the measure $\mu \llcorner A$: $2^X \to [0,\infty]$ given by*
$$(\mu \llcorner A)(B) := \mu(A \cap B)$$
is a measure.

1.1. **Example: Hausdorff measure.** Let (X,d) be a metric space.

Definition 1.7. The s-dimensional *Hausdorff measure*, $s > 0$ is defined as follows.

Let $\delta \in (0,\infty]$, then for any $A \subset X$ we define
$$\mathcal{H}^s_\delta(A) := \alpha(s) \inf \left\{ \sum_{k=1}^\infty r_k^s : \quad A \subset \bigcup_{k=1}^\infty B(x_k, r_k), \quad r_k \in (0,\delta) \right\}.$$
Here $B(x_k, r_k)$ are open balls with radius r centered at x_k, i.e.
$$B(x_k, r_k) := \{y \in X : d(x_k, y) < r_k\}.$$
Moreover[2]
$$\alpha(s) := \frac{\pi^{\frac{s}{2}}}{\Gamma(\frac{s}{2}+1)}.$$
where Γ is the Γ-function.

Now observe that $\delta \mapsto \mathcal{H}^s_\delta(A)$ is monotonce decreasing. So we can write
$$\mathcal{H}^s(A) := \lim_{\delta \to 0^+} \mathcal{H}^s_\delta(A) \equiv \sup_{\delta > 0} \mathcal{H}^s_\delta(A) \in [0,\infty].$$

Often one writes $\mathcal{H}^0(A) := \#A$, the *counting measure*.

$\mathcal{H}^s_\infty$ is called the *Hausdorff content*.

[2]Warning: Some authors set $\alpha(s) := 1$. The main reason to not do that is so that $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$

Remark 1.8.
- Observe that while $\mathcal{H}^s_\delta(A) < \infty$ whenever $s > 0$, $\delta > 0$ and A is any bounded set, as $\delta \to 0$ $\mathcal{H}^s(A)$ will be infinite whenever s is smaller than the "dimension of A" (a notion we will define more carefully below).

Lemma 1.9. *$\mathcal{H}^s$ is a measure in $\mathbb{R}^n$.*

Proof. One can show similar to the argument for m^* that $\mathcal{H}^s_\delta(\cdot)$ is a measure for each $\delta > 0$.

We clearly have $\mathcal{H}^s(\emptyset) = 0$. Moreover, since $\mathcal{H}^s_\delta$ is a measure for any $\delta > 0$, we have for any $A \subset \bigcup_{k=1}^\infty A_k$,

$$\mathcal{H}^s_\delta(A) \leq \sum_{k=1}^\infty \mathcal{H}^s_\delta(A_k) \leq \sum_{k=1}^\infty \mathcal{H}^s(A_k).$$

Taking the supremum over δ in this inequality we have σ-additivity for $\mathcal{H}^s$.

$$\mathcal{H}^s(A) \leq \sum_{k=1}^\infty \mathcal{H}^s(A_k).$$

□

Exercise 1.10. *Show that*

$$\mathcal{H}^0_\delta(\mathbb{Q}) \xrightarrow{\delta \to 0} \infty.$$

Lemma 1.11. *$\mathcal{H}^s$ is a metric outer measure that means that if $A, B \subset (X, d)$ satisfy*

$$d(A,B) := \inf_{a \in A, b \in B} d(a,b) > 0.$$

then

$$\mathcal{H}^s(A \cup B) = \mathcal{H}^s(A) + \mathcal{H}^s(B).$$

Proof. This is relatively easy to see. Take $\delta < \frac{d(A,B)}{3}$, then since any covering $B(x,r)$ with $r < \delta$ cannot contain points of both A or B at the same time, we have that $\mathcal{H}^s_\delta$ is a metric outer measure. Taking $\delta \to 0^+$ we obtain that $\mathcal{H}^s$ is a metric outer measure. □

Remark 1.12. One can, and we will in Corollary 1.78, show that the n-dimensional Hausdorff measure in $\mathbb{R}^n$ coincides with the Lebesgue measure $\mathcal{L}^n$, i.e.

$$\mathcal{L}^n(A) = \mathcal{H}^n(A).$$

Exercise 1.13. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be (uniformly) Lipschitz continuous that is*

$$|f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in \mathbb{R}^n$$

Then for any set Ω and any $s \geq 0$,

$$\mathcal{H}^s(f(\Omega)) \leq C(L)\mathcal{H}^s(\Omega)$$

where $C(L)$ is a constant only depending on L.

Exercise 1.14. *Show that*

- $\mathcal{H}^1 = \mathcal{L}^1$ *in* $\mathbb{R}$

- *$\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$ for all $\lambda > 0$, where $\lambda A = \{\lambda x : \quad x \in A\}$.*
- *$\mathcal{H}^s(LA) = \mathcal{H}^s(A)$ whenever $L : \mathbb{R}^n \to \mathbb{R}^n$ is an affine isometry, i.e. if $Lx = Ax + b$ for $A \in O(n)$ and $b \in \mathbb{R}^n$ constant.*

Exercise 1.15. *Let $U \subset \mathbb{R}^n$ be any non-empty open set. Then $\mathcal{H}^s(U) = \infty$ for all $s < n$.*

Exercise 1.16 (translation and rotation invariant)**.** *Let $A \subset \mathbb{R}^n$ and $s \in (0, \infty)$. Show the following*

(1) If $p \in \mathbb{R}^n$ then $\mathcal{H}^s(p + A) = \mathcal{H}^s(A)$.
(2) If $O \in O(n)$ (i.e. $O \in \mathbb{R}^{n \times n}$ and $O^t O = I$) then $\mathcal{H}^s(OA) = \mathcal{H}^s(A)$.
(3) If $A \subset \mathbb{R}^\ell \times \{0\}$ for $0 < \ell < n$ and $\pi : (x_1, \ldots, x_n) := (x_1, \ldots, x_\ell)$ is the projection from $\mathbb{R}^n = \mathbb{R}^\ell \times \mathbb{R}^{n-\ell}$ to $\mathbb{R}^\ell$, then $\mathcal{H}^s_{\mathbb{R}^n}(A) = \mathcal{H}^s_{\mathbb{R}^\ell}(\pi(A))$.

Lemma 1.17. *Let $0 \leq s < t < \infty$.*

(1) If $\mathcal{H}^s(A) < \infty$ then $\mathcal{H}^t(A) = 0$
(2) If $\mathcal{H}^t(A) > 0$ then $\mathcal{H}^s(A) = \infty$.

Proof. Indeed, whenever $r_k \leq \delta$ and $(B(x_k, r_k))_{k \in \mathbb{N}}$ cover A we have

$$\mathcal{H}^t_\delta(A) \leq \alpha(t) \sum_{k=1}^\infty r_k^t \leq \alpha(t) \delta^{t-s} \sum_{k=1}^\infty r_k^s.$$

Taking the infimum over any such covering $B(x_k, r_k)$ of A we find

$$\mathcal{H}^t_\delta(A) \leq \frac{\alpha(t)}{\alpha(s)} \delta^{t-s} \mathcal{H}^s_\delta(A).$$

Taking $\lim_{\delta \to 0}$ on both sides we obtain

$$\mathcal{H}^t(A) \leq \frac{\alpha(t)}{\alpha(s)} 0 \cdot \mathcal{H}^s(A).$$

This implies that if $\mathcal{H}^t(A) > 0$ then necessarily $\mathcal{H}^s(A) = \infty$, and if $\mathcal{H}^s(A) < \infty$ then $\mathcal{H}^t(A) = 0$. $\square$

Example 1.18. If $k \in \mathbb{N}$ it is conceivable that $\mathcal{H}^k$ measures something of "dimension k". For example assume that $C = [0,1]^2 \times \{0\} \subset \mathbb{R}^3$ is a 2D-square of sidelength 1. We need $\approx \frac{1}{\delta^2}$ many balls to cover C. Then

$$\mathcal{H}^s_\delta(C) \leq \alpha(s) \frac{1}{\delta^2} \delta^s.$$

So if $s > 2$ we see that $\mathcal{H}^s(C) \leq \lim_{\delta \to 0} \delta^{s-2} = 0$. That is C has no s-volume for $s > 2$.

For $s = 2$ one can argue that covering uniformly by balls of radius δ is optimal and thus we have

$$0 < \mathcal{H}^2(C) < \infty.$$

In particular $\mathcal{H}^s(C) = \infty$ for any $s < 2$.

(this argument is easy to generalize to a ℓ-dimensional manifold in $\mathbb{R}^N$)

Indeed, with the Hausdorff measure we can define a dimension

Definition 1.19. The *Hausdorff dimension* is defined as

$$\dim_{\mathcal{H}} A := \inf \{ s \geq 0 : \quad \mathcal{H}^s(A) = 0 \} .$$

If $\mathcal{H}^s(A) > 0$ for all $s > 0$ then $\dim_{\mathcal{H}}(E) := \infty$.

Lemma 1.20. *Let C be a set in a metric space and let $s \geq 0$*

(1) *If $\mathcal{H}^s(C) = 0$ then $\dim_{\mathcal{H}}(E) \leq s$.*
(2) *If $\mathcal{H}^s(C) > 0$ then $\dim_{\mathcal{H}}(E) \geq s$.*
(3) *If $0 < \mathcal{H}^s(C) < \infty$ then $\dim_{\mathcal{H}}(E) = s$.*
(4) *If $\mathcal{H}^s_\infty(C) > 0$ and $\mathcal{H}^s(C) < \infty$ then $\dim_{\mathcal{H}}(E) = s$.*

Proof. This follows from Lemma 1.17 and the definition of Hausdorff measure.

(1) follows from the definition of the Hausdorff measure as infimum. then $\dim_{\mathcal{H}}(E) \leq s$.
(2) If $\mathcal{H}^s(C) > 0$ then by Lemma 1.17 $\mathcal{H}^t(C) = \infty$ for all $t < s$. Again from the definition it is clear that $\dim_{\mathcal{H}}(E) \geq s$.
(3) This is a consequence of the two above statements.
(4) Follows from the statement before since $\mathcal{H}^s_\infty(C) \leq \mathcal{H}^s(C)$

□

Exercise 1.21 (Hausdorff dimension under Lipschitz and Hölder maps)**.** *Let (X, d_x) and (Y, d_Y) be two metric spaces and let $f : X \to Y$. Assume that $A \subset X$ has Hausdorff-dimension $\dim_{\mathcal{H}}(A) = s$.*

(1) *If f is uniformly Lipschitz continuous, i.e. for some $L > 0$,*
$$d_Y(f(x), f(y)) \leq L\, d(x, y) \quad \forall x, y \in X$$
then $\dim_{\mathcal{H}}(f(A)) \leq s$.
(2) *Give an example where $\dim_{\mathcal{H}}(A) < s$*
(3) *Assume f is uniformly Hölder continuous, i.e. for some $L > 0$ and $\alpha > 0$*
$$d_Y(f(x), f(y)) \leq L\, d(x, y)^\alpha \quad \forall x, y \in X$$
What can we say about the Hausdorff dimension of $f(A) \subset Y$?

Cf Exercise 1.13.

Example 1.22. The Cantorset is defined as follows.

$$C_0 := [0, 1]$$

Let $C_0 := [0, 1]$. In the k-th step we construct C_k by removing of each interval the open middle interval of size 3^{-n}. For exmaple

$$C_1 := [0, \frac{1}{3}] \cup [\frac{1}{3}, 1].$$

FIGURE 1.2. The cantor set

See Figure 1.2.

Set $C := \bigcap_{k=1}^{\infty} C_k$. Observe that C is closed and bounded, so compact.

Lemma 1.23. $\dim_{\mathcal{H}}(C) = \frac{\log 2}{\log 3}$.

Proof. For each $k \in \mathbb{N}$ we have $C \subset C_k$. Observe that C_k consists of 2^k disjoint intervals each of diameter 3^{-k} (i.e. radius $\frac{1}{2}3^{-k}$). Thus for any $\delta > 0$ and for any $k \gg 1$ so that $\frac{1}{2}3^{-k} < \delta$ we have

$$\mathcal{H}^s_\delta(C) \leq \alpha(s) \sum_{\ell=1}^{2^k} (\frac{1}{2}3^{-k})^s = 2^{-s} \left(\frac{2}{3^s}\right)^k \xrightarrow{k\to\infty} \alpha(s) \begin{cases} 2^{-s} & s = \frac{\log 2}{\log 3} \\ 0 & s > \frac{\log 2}{\log 3} \\ \infty & s > \frac{\log 2}{\log 3} \end{cases}$$

In particular we have

$$\mathcal{H}^s(C) = 0 \quad \forall s > \frac{\log 2}{\log 3}.$$

So from the definition of the Hausdorff dimension we get

$$\dim_{\mathcal{H}} C \leq \frac{\log 2}{\log 3}.$$

Now we need to show the <u>other direction</u>. From now on set $s := \frac{\log 2}{\log 3}$. Let $(B(x_i, r_i))_{i=1}^{\infty}$ be any covering of C. We claim that

$$\sum_{i=1}^{\infty} r_i^s \geq \frac{1}{2^s 4}. \tag{1.3}$$

Once we have (1.3) we are done, because (1.3) implies

$$\mathcal{H}^s\infty(C) \geq \frac{1}{2^s 4}.$$

In particular (recall that $s = \frac{\log 2}{\log 3}$) we have $\infty > \mathcal{H}^s(C) \geq \mathcal{H}^s_\infty(C) > 0$.

Let us make some notation. Denote by A_k the intervals of C_k, i.e. A_k consists of pairwise disjoint, closed intervals in $\mathbb{R}$ such that $C_k = \bigcup_{I \in A_k} I$. E.g.

$$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \quad A_1 = \{[0, \frac{1}{3}], [\frac{2}{3}, 1]\}.$$

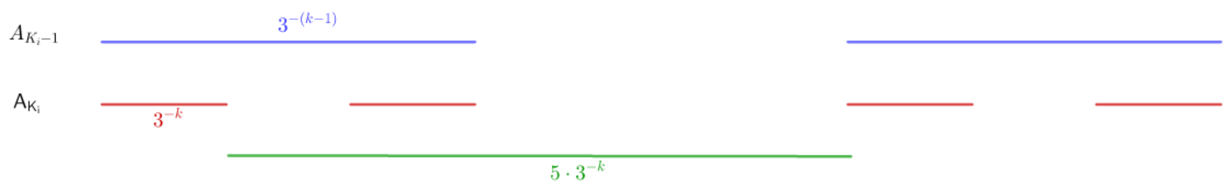


FIGURE 1.3. If a ball intersects three intervals of A_{K_i} its *diameter* is at least $5 \cdot 3^{-K_i}$

Proof of (1.3) Since C is compact, we may assume that there finitely many, w.l.o.g. the first N balls $(B(x_i, r_i))_{i=1}^N$ already cover C. We may assume that each $r_i < \frac{1}{2}$, otherwise (1.3) is obvious.

Fix $i \in \{1, \ldots, N\}$.

Let $K_i \in \mathbb{N} \cup \{0\}$ so that

$$2r_i \in [3^{-K_i-1}, 3^{-K_i}).$$

Now we consider the construction step C_{K_i}. Each ball $B(x_i, r_i)$ has nonempty intersection with at most 2 intervals of C_{K_i}. Indeed, otherwise its *diameter* would be at least $5 \cdot 3^{-K_i}$, see Figure 1.3.

But then $B(x_i, r_i)$ has nonempty intersection with at most $2 \cdot 2^{j-K_i}$ intervals of C_j for any $j \geq K_i$. Since $s = \frac{\log 2}{\log 3}$ we have

$$2 \cdot 2^{j-K_i} = 2^{j+1} 2^{-K_i} = 2^{j+1} 3^{-K_i s} \leq 2^{j+1} 3^s (2r_i)^s = 2^{j+2} (2r_i)^s.$$

Set now $K := \max_{\{i=1,\ldots,N\}} K_i$.

Then for any $i \in \{1, \ldots, N\}$ each of the balls $B(x_i, r_i)$ has nonempty intersection with at most $2^{K+2}(2r_i)^s$ many intervals of A_K.

So if we set Γ_i to be the number of intervals in A_K that intersect $B_{r_i}(x_i)$ we have $\Gamma_i \leq 2^{K+2}(2r_i)^s$ and thus

$$\begin{aligned} \sum_{i=1}^N \Gamma_i (3^{-K})^s \leq & \sum_{i=1}^N \underbrace{(3^{-K})^s}_{=2^{-K}} 2^{K+2} (2r_i)^s \\ =& 4\; 2^s \sum_{i=1}^N (r_i)^s \end{aligned} \tag{1.4}$$

Now for each $x \in C$ there is exactly one interval I in A_K such that $x \in I$. Since $(B_{r_i}(x_i))_{i=1}^N$ covers all of C we have the following: for each interval I in A_K there exists some $i \in \{1, \ldots, N\}$ such that $B_{r_i}(x_i) \cap I \neq \emptyset$. That is,

$$\sum_{i=1}^N \Gamma_i \geq \text{number of intervals in } A_K = 2^K.$$

FIGURE 1.4. The fat cantor set for $a = \frac{1}{4}$, see Example 1.24

Thus,

$$\sum_{i=1}^{N} \Gamma_i 3^{-Ks} \geq 2^K 3^{-Ks} = 1. \tag{1.5}$$

Together, (1.4) and (1.5) imply (1.3). □

Example 1.24. The Smith–Volterra–Cantor set, aka fat cantor set is defined as follows.

Let $C_0 := [0, 1]$. In the k-th step we construct C_k by removing of each interval the open middle interval of size a^n. That is

$$C_1 = [0, \frac{1-a}{2}] \cup [\frac{1+a}{2}, 1].$$

$$C_2 = [0, \frac{1-a}{4} - \frac{a^2}{2}] \cup [\frac{1-a}{4} + \frac{a^2}{2}, \frac{1-a}{2}] \cup [\frac{1+a}{2}, \frac{1+\frac{1+a}{2}}{2} - \frac{a^2}{2}] \cup [\frac{1+\frac{1+a}{2}}{2} + \frac{a^2}{2}, 1].$$

Cf. Figure 1.4.

Set $C := \bigcap_{k=1}^{\infty} C_k$. For $a = \frac{1}{3}$ this is the typical *Cantor set*. For $a = \frac{1}{4}$ this is the *Fat Cantor set*.

Exercise 1.25. *The fat Cantor above set has positive $\mathcal{H}^1$-measure.*

1.2. **Measurable sets.** As we have discussed, our definition of measure does not include the "natural" condition that $\mu(B) = \mu(B \cap A) + \mu(B \backslash A)$ for all $A, B \subset X$ – because this "natural" condition leads to incompatibility such as the Banach-Tarski Paradoxon.

So we will denote the class of sets $A \subset 2^X$ where we have the above "natural" condition as the σ-algebra of measurable sets.

Definition 1.26 (Carathéodory)**.** Let μ be a measure on X.

$A \subset X$ is called *μ-measurable* if

$$\mu(B) = \mu(A \cap B) + \mu(B \backslash A) \quad \text{for any } B \subset X$$

Remark 1.27. By additivity of the measure, measurability is equivalent to

$$\mu(B) \geq \mu(A \cap B) + \mu(B \backslash A) \quad \text{for any } B \subset X$$

Exercise 1.28. *Let $X \neq \emptyset$ be any set*

- *and assume $\mu(\emptyset) = 0$ and $\mu(A) = 1$ for any $A \neq \emptyset$. Then A is μ-measurable if and only if $A = \emptyset$ or $A = X$.*
- *If $\nu = \#$ the counting measure then any set A is ν-measurable.*

Clearly, whatever choice of measure we have, $\emptyset$ and X are measurable sets. We also have

$$(A_i)_{i=1}^N \text{ are measurable} \quad \Rightarrow \quad \bigcup_{i=1}^N A_i \text{ is measurable} \tag{1.6}$$

Proof of (1.6). We proof this by induction. Clearly this holds for $N = 1$. So to conclude (1.6) we only need to show:

If A_1, A_2 are μ-measurable, then so is $A_1 \cup A_2$.

So assume A_1 and A_2 are μ-measurable and $B \subset X$.

$$\begin{aligned}
\mu(B) =&\mu(B \setminus A_1) + \mu(B \cap A_1)\\
=&\mu\left((B \setminus A_1) \cap A_2\right) + \mu\left((B \setminus A_1) \setminus A_2\right)\\
&+ \mu\left((B \cap A_1) \cap A_2\right) + \mu\left((B \cap A_1) \setminus A_2\right)\\
\geq&\mu\left(B \setminus (A_1 \cup A_2)\right) + \mu(B \cap (A_1 \cup A_2))
\end{aligned}$$

In the last step we have used that

$$\mu\left((B \setminus A_1) \cap A_2\right) + \mu\left((B \cap A_1) \cap A_2\right) + \mu\left((B \cap A_1) \setminus A_2\right) \geq \mu\left(B \setminus (A_1 \cup A_2)\right),$$

by sublinearity and the fact that

$$B \setminus (A_1 \cup A_2) = \left((B \setminus A_1) \cap A_2\right) \cup \left((B \cap A_1) \cap A_2\right) \cup \left((B \cap A_1) \setminus A_2\right).$$

By Remark 1.27 we have that $(A_1 \cup A_2)$ is also measurable. □

We have much more than that:

Lemma 1.29. *Let X be a set and μ be a measure on X.*

The collection $\mathcal{A} \subset 2^X$ of μ-measurable functions

$$\mathcal{A} := \{A \subset X : \quad A \text{ is } \mu\text{-measurable}\}$$

is a σ-algebra, that is

(1) *$X \in \mathcal{A}$*
(2) *$A \in \mathcal{A}$ implies that $X \backslash A \in \mathcal{A}$*
(3) *If $(A_i)_{i=1}^\infty \subset \mathcal{A}$ then $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$.*[3]

In particular

[3] This is the σ in σ-algebra, σ means for countably many. If we only had for any $N \in \mathbb{N}$: $(A_i)_{i=1}^N \subset \mathcal{A}$ then $\bigcup_{i=1}^N A_i \in \mathcal{A}$, $\mathcal{A}$ would be merely an Algebra (no σ!)

- $\emptyset \in \mathcal{A}$
- *if* $(A_i)_{i=1}^\infty \subset \mathcal{A}$ *then* $\bigcap_{i=1}^\infty A_i \in \mathcal{A}$

Proof. (1) For any $B \subset X$: since $B \cap X = B$ and $B \backslash X = \emptyset$ we have

$$\mu(B) = \mu(B) + \mu(\emptyset) = \mu(B \cap X) + \mu(B \setminus X).$$

(2) Assume that $A \in \mathcal{A}$. Set $\tilde{A} := X \backslash A$. For any $B \subset X$ we have

$$\tilde{A} \cap B = (X \backslash A) \cap B = B \backslash A,$$

and

$$B \setminus \tilde{A} = B \setminus (X \backslash A) = B \cap A.$$

Since A is measurable we then have

$$\mu(B \cap \tilde{A}) + \mu(B \setminus \tilde{A}) = \mu(B \setminus A) + \mu(B \cap A) = \mu(B).$$

(3) Let $(A_i)_{i \in \mathbb{N}} \subset \mathcal{A}$. Set $A := \bigcup_{i=1}^\infty A_i$.

Without loss of generality we have that $A_i \cap A_j = \emptyset$ for $i \neq j$. Indeed, otherwise we set $\tilde{A}_1 := A_1$ and $\tilde{A}_k := A_k \backslash \bigcup_{i=1}^{k-1} A_i$. By the previously proven properties and (1.6) each $\tilde{A}_k$ belongs to $\mathcal{A}$ and we have $A = \bigcup_{k=1}^\infty \tilde{A}_k$ – so we could work with $\tilde{A}_k$ instead of A_k.

We have by measurability of each A_k and since A_N and $\bigcup_{k=1}^{N-1} A_k$ are disjoint,

$$\begin{aligned}\mu(B \cap \bigcup_{k=1}^N A_k) =& \mu(B \cap \left(\bigcup_{k=1}^N A_k\right) \cap A_N) + \mu(\left(B \cap \bigcup_{k=1}^N A_k\right) \setminus A_N) \\ =& \mu(B \cap A_N) + \mu(B \cap \bigcup_{k=1}^{N-1} A_k)\end{aligned}$$

Repeating this computation $N-1$ times we obtain

$$\mu(B \cap \bigcup_{k=1}^N A_k) = \sum_{k=1}^N \mu(B \cap A_k). \tag{1.7}$$

By (1.6) and the monotonicity of μ, Remark 1.2, we then have

$$\mu(B) = \mu(B \cap \bigcup_{k=1}^N A_k) + \mu(B \setminus \bigcup_{k=1}^N A_k) \geq \sum_{k=1}^N \mu(B \cap A_k) + \mu(B \backslash \bigcup_{k=1}^\infty A_k)$$

This holds for any N, so we obtain

$$\mu(B) \geq \sum_{k=1}^\infty \mu(B \cap A_k) + \mu(B \backslash \bigcup_{k=1}^\infty A_k)$$

By the σ-subadditivity of μ we then have

$$\mu(B) \geq \mu(B \cap \bigcup_{k=1}^\infty A_k) + \mu(B \backslash \bigcup_{k=1}^\infty A_k)$$

In view of Remark 1.27 this implies measurability of $\bigcup_{k=1}^\infty A_k$.

□

Definition 1.30. Let $\mathcal{C} \subset 2^X$ any nonempty family of subsets of X, then

$$\sigma(\mathcal{C})$$

denotes the *σ-Algebra generated by $\mathcal{C}$*, namely the smallest σ-algebra containing $\mathcal{C}$.

Exercise 1.31.
- *$\{\emptyset, X\}$ is a σ-algebra of X*
- *2^X is a σ-algebra of X*
- *Let (X, d) be a metric space. Denote $\mathcal{O} \subset 2^X$ the family of all open sets.*

  Let $\mathcal{F}$ be the family of σ-Algebras that contain all open sets. That is, $\mathcal{A} \subset 2^X$ belongs to $\mathcal{F}$ if and only if $\mathcal{A}$ is a σ-Algebra, and any open set $O \in \mathcal{O}$ belongs to $\mathcal{A}$, i.e. $O \in \mathcal{A}$.

  Define
  $$\mathcal{B} := \bigcap \{\mathcal{A} : \quad \mathcal{A} \in \mathcal{F}\}.$$
  Show that (a) $\mathcal{F}$ is nonempty, (b) $\mathcal{B}$ is a σ-algebra and (c) $\mathcal{B}$ is the smallest σ-Algebra containing all open sets, i.e. show that $\mathcal{B} = \sigma(\mathcal{O})$.

  $\mathcal{B}$ is called the Borel σ-Algebra *and a set $B \in \mathcal{B}$ is called a* Borel set.

Definition 1.32. If $\mu : 2^X \to [0, \infty]$ is a measure on X, and Σ is the σ-algebra of μ-measurable sets, then once calls (X, Σ, μ) a *measure space*.

Some author choose to define measures only on their σ-algebra Σ of measurable sets, and call our definition of a measure an outer measure.

There is a reason for restricting μ only to act on measurable sets – a measure μ acts in a very intuitive way on its measurable sets!

Theorem 1.33. *Let (X, Σ, μ) be a measure space.*

Let $(A_k)_{k \in \mathbb{N}} \subset \Sigma$ (i.e. each A_k is measurable). Then we have

(1) *If $A_k \cap A_\ell = \emptyset$ for $k \neq \ell$ we have*

(σ-Additivity)
$$\mu(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} \mu(A_k)$$

(2) *If $A_1 \subset A_2 \subset \ldots \subset A_k \subset \ldots$ then*
$$\mu(\bigcup_{k=1}^{\infty} A_k) = \lim_{k \to \infty} \mu(A_k).$$

(3) *If $\mu(A_1) < \infty$ and $A_1 \supset A_2 \supset \ldots \supset A_k \supset \ldots$ then*
$$\mu(\bigcap_{k=1}^{\infty} A_k) = \lim_{k \to \infty} \mu(A_k).$$

Proof. (1) Above in (1.7) we computed (take $B = X$) that for finitely many pairwise disjoint sets

$$\mu(\bigcup_{k=1}^{N} A_k) = \sum_{k=1}^{N} \mu(A_k).$$

By monotonicity

$$\mu(\bigcup_{k=1}^{\infty} A_k) \geq \mu(\bigcup_{k=1}^{N} A_k) = \sum_{k=1}^{N} \mu(A_k)$$

Taking $N \to \infty$ we find

$$\mu(\bigcup_{k=1}^{\infty} A_k) \geq \sum_{k=1}^{\infty} \mu(A_k).$$

By σ-subadditivity,

$$\mu(\bigcup_{k=1}^{\infty} A_k) \geq \sum_{k=1}^{\infty} \mu(A_k) \geq \mu(\bigcup_{k=1}^{\infty} A_k).$$

The number on left- and right-hand side are the same so we have

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(\bigcup_{k=1}^{\infty} A_k).$$

(2) Let $\tilde{A}_1 := A_1$ and $\tilde{A}_k := A_k \backslash A_{k-1}$. Then $(\tilde{A}_k)_{k=1}^{\infty}$ is pairwise disjoint, and each $\tilde{A}_k$ is measurable. So

$$\begin{aligned}\mu(\bigcup_{k=1}^{\infty} A_k) =& \mu(\bigcup_{k=1}^{\infty} \tilde{A}_k) = \sum_{k=1}^{\infty} \mu(\tilde{A}_k)\\ =& \lim_{N\to\infty} \sum_{k=1}^{N} \mu(\tilde{A}_k)\\ =& \lim_{N\to\infty} \mu(\bigcup_{k=1}^{N} \tilde{A}_k)\\ =& \lim_{N\to\infty} \mu(A_N).\end{aligned}$$

(3) Set $\tilde{A}_k := A_1 \backslash A_k$, $k \in \mathbb{N}$. Then $\emptyset = \tilde{A}_1 \subset \tilde{A}_2 \subset \ldots$. Moreover we have

$$\mu(A_1) = \mu(\tilde{A}_k) + \mu(A_k), \quad k \in \mathbb{N}.$$

By the above argument (observe that $\mu(A_k) \le \mu(A_1) < \infty$)

$$\begin{aligned}\mu(A_1) - \lim_{k\to\infty} \mu(A_k) = &\lim_{k\to\infty} \mu(\tilde{A}_k)\\ =&\mu(\bigcup_{k=1}^{\infty} \tilde{A}_k)\\ =&\mu(A_1 \backslash \bigcap_{k=1}^{\infty} A_k)\\ =&\mu(A_1) - \mu(\bigcap_{k=1}^{\infty} A_k).\end{aligned}$$

Since $\mu(A_1) < \infty$ we can conclude.

□

Example 1.34.
- There is no way we can assume (or should hope for) that for uncountable unions we have (even sub-)additivity: For example $\mathbb{R} = \bigcup_{x\in\mathbb{R}}\{x\}$. The Lebesgue measure would satisfy $\mathcal{L}^1\{x\} = 0$ for all x, so
$$\mathcal{L}^1(\mathbb{R}) \neq \lim_{k\to\infty} \mathcal{L}^1\{x\} = 0.$$
- The assumption $\mu(A_1) < \infty$ in Theorem 1.33(3) is necessary.
  Let $X = \mathbb{N}$ and μ the counting measure. Set for $k \in \mathbb{N}$
$$A_k = \{k, k+1, k+2, \ldots\}.$$
  Then $A_k \supset A_{k+1}$, but $\mu(A_k) = \infty$ for all $k \in \mathbb{N}$. However $\bigcap_{k\in\mathbb{N}} A_k = \emptyset$, so
$$0 = \mu(\bigcap_{k\in\mathbb{N}} A_k) \neq \lim_{k\to\infty} \mu(A_k) = \infty.$$

Lastly, from Advanced Calculus we are aware of sets of measure zero (there it was that Riemann-integrable functions are continuous outside a set of *Lebesgue*-measure zero)

Definition 1.35 (Zero sets)**.** A set $N \subset X$ is called a set of *μ-measure zero* if $\mu(N) = 0$. We also say N is a *μ-zeroset*.

A property $P(x)$ holds *μ-a.e.* in X if $P(x)$ holds in $X\backslash N$ where N is a μ-zeroset.

Theorem 1.36. *Let $N \subset X$ be a μ-zero set. Then N is measurable.*

Proof. Let $B \subset X$, by monotonicity we have $\mu(B \cap N) \le \mu(N) = 0$. Moreover $\mu(B\backslash N) \le \mu(B)$. So we have
$$\mu(B \cap N) + \mu(B\backslash N) \le \mu(N) + \mu(B) = \mu(B).$$
This implies measurability. □

Exercise 1.37. *Let X be a set and μ be a measure on X.*

(1) Let $(N_i)_{i\in\mathbb{N}}$ be sets of μ-measure zero. Show that $\bigcup_{i\in\mathbb{N}} N_i$ is a set of μ-measure zero.

(2) *Let N be a μ-zeroset. Show that any $A \subset N$ is a μ-zeroset.*
(3) *Show that is a property $P(x)$ holds for μ-a.e. x and a property $Q(y)$ holds for μ-a.e. y then $Q(x)$ and $P(x)$ hold (simultaneously) for a.e. x.*

1.3. **Construction of Measures: Carathéodory-Hahn Extension Theorem.** While we already have constructed in (1.2) the Lebesgue (outer) measure, we are still interested in finding a more axiomatic approach to construct the Lebesgue measure.

The idea is to define a *pre-measure* on some sets (like cubes!) and build a measure out of that.

Definition 1.38 (Algebra)**.** Let X be a set and $\mathcal{A} \subset 2^X$. $\mathcal{A}$ is called an *algebra* if

(1) $X \in \mathcal{A}$
(2) $A \in \mathcal{A}$ implies that $X \backslash A \in \mathcal{A}$
(3) If $A_1, A_2 \in \mathcal{A}$ then $A_1 \cup A_2 \in \mathcal{A}$.

Definition 1.39 (Pre-measure)**.** Let X be a set, $\mathcal{A} \subset 2^X$ an algebra (not necessarily a σ-algebra!). A map $\lambda : \mathcal{A} \to [0, \infty]$ is a *pre-measure*, if

(1) $\lambda(\emptyset) = 0$
(2) $\lambda(A) = \sum_{k=1}^{\infty} \lambda(A_k)$ for any $A \in \mathcal{A}$ such that $A = \bigcup_{k=1}^{\infty} A_k$ for some parwise disjoint $(A_k)_{k=1}^{\infty} \subset \mathcal{A}$.

A pre-measure is called *σ-finite*, if $X = \bigcup_{k=1}^{\infty} S_k$ with $S_k \in \mathcal{A}$ and $\lambda(S_k) < \infty$ for each k.

Exercise 1.40. *Let $\lambda : \mathcal{A} \to [0, \infty]$ be a premeasure*

(1) *Assume $A \subset B$ with $A, B \in \mathcal{A}$ then $\lambda(A) \leq \lambda(B)$.*

Example 1.41. An interval is the set (a, b) or $[a, b]$ or $(a, b]$ or $[a, b)$, where $-\infty < a \leq b < \infty$ (we allow $\pm\infty$ if the set is open).

A *block* Q in $\mathbb{R}^n$ is the cartesian product of n intervals $Q = I_1 \times I_2 \times \ldots \times I_n$.

A *figure* is made out of finitely many blocks

$$\mathcal{A} := \left\{ A \subset \mathbb{R}^n : A = \bigcup_{i=1}^{N} Q_i \quad \text{some } N \in \mathbb{N},\ Q_i \text{ blocks with pairwise disjoint interior} \right\}.$$

Exercise: $\mathcal{A}$ is an Algebra

The volume of a block $Q = I_1 \times \ldots \times I_n$ is given by the n-dimensional volume, i.e.

$$\mathrm{vol}(Q) = \Pi_{i=1}^{n} |I_i|,$$

where as usual, $|[a, b]| = |(a, b)| = |[a, b)| = |(a, b]| := |b - a|$.

Indeed, vol defines now a premeasure on $\mathcal{A}$:

Whenever $A \in \mathcal{A}$, i.e. $A = \bigcup_{i=1}^{N} Q_i$ for pairwise disjoint blocks Q_i then

$$\mathrm{vol}(A) := \sum_{i=1}^{N} \mathrm{vol}(Q_i).$$

One can check that this is independent of the specific choice of Q_i. I.e. if also $A = \bigcup_{j=1}^{\tilde{N}} \tilde{Q}_j$ for another cobination of pairwise disjoint blocks $\tilde{Q}_j$ then

$$\sum_{i=1}^{N} \mathrm{vol}(Q_i) = \sum_{j=1}^{\tilde{N}} \mathrm{vol}(\tilde{Q}_j).$$

Indeed, now vol defines a *pre-measure* on $\mathcal{A}$. vol is also σ-finite, simply take $S_k := [-k,k]^n$.

The important thing is that a premeasure extends (more or less uniquely) into a real measure. This is called *Carathéodory–Hahn extension*.

Theorem 1.42 (Carathéodory–Hahn extension)**.** *Let X be a set, and $\mathcal{A} \subset 2^X$ an algebra with a premeasure $\lambda : \mathcal{A} \to [0,\infty]$.*

For $A \subset X$ the Carathéodory–Hahn extension *$\mu : 2^X \to [0,\infty]$ is defined as*

$$\mu(A) := \inf \left\{ \sum_{k=1}^{\infty} \lambda(A_k) : \quad A \subset \bigcup_{k=1}^{\infty} A_k, \quad A_k \in \mathcal{A} \right\}$$

Then

(1) $\mu : 2^X \to [0,\infty]$ is a measure on X,
(2) $\mu(A) = \lambda(A)$ for all $A \in \mathcal{A}$,
(3) Any $A \in \mathcal{A}$ is μ-measurable.

Compare this to the definition of the Lebesgue measure, (1.2).

Proof of Theorem 1.42. (1) Clearly $\mu : 2^X \to [0,\infty]$ is well-defined (observe that $X \in \mathcal{A}$, so that $\mu(B) \leq \mu(X)$ for all $B \subset X$, and $\mu(\emptyset) = 0$.

Now let $B \subset \bigcup_{k=1}^{\infty} B_k$. Take any $\varepsilon > 0$. By definition of the infimum, for each B_k there exist some $(A_{k;\ell})_{\ell=1}^{\infty} \subset \mathcal{A}$ such that $B_k \subset \bigcup_{\ell=1}^{\infty} A_{k;\ell}$ and

$$\sum_{\ell=1}^{\infty} \lambda(A_{k;\ell}) \leq \mu(B_k) + 2^{-k}\varepsilon.$$

Clearly, $A \subset \bigcup_{k,\ell=1}^{\infty} A_{k;\ell}$ so we have

$$\begin{aligned}\mu(A) &\leq \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} \lambda(A_k;\ell)\\ &\leq \sum_{k=1}^{\infty}\Big(\mu(B_k)+2^{-k}\varepsilon\Big)\\ &\leq \left(\sum_{k=1}^{\infty}\mu(B_k)\right)+\varepsilon.\end{aligned}$$

This holds for any $\varepsilon > 0$, so letting $\varepsilon \to 0$ we find

$$\mu(A) \leq \left(\sum_{k=1}^{\infty}\mu(B_k)\right)+\varepsilon.$$

Thus, μ is σ-subaddive, and thus μ is a measure.

(2) For $A \in \mathcal{A}$ we clearly have $\mu(A) \leq \lambda(A)$.

Now we show $\lambda(A) \leq \mu(A)$. We may assume that $\mu(A) < \infty$ otherwise there is nothing to show.

Take any $(A_k)_{k=1}^{\infty} \subset \mathcal{A}$ such that $\bigcup_{k=1}^{\infty} A_k \supset A$. By the usual argument we may assume (without loosing that $A_k \in \mathcal{A}$) that $A_k \cap A_j = \emptyset$ for all $k \neq j$. Set $\tilde{A}_k := A_k \cap A$, $k \in \mathbb{N}$. Then $(\tilde{A}_k)_{k=1}^{\infty} \in \mathcal{A}$ are pairwise disjoint, so since λ is premeasure

$$\lambda(A) = \sum_{k=1}^{\infty}\lambda(\tilde{A}_k) \leq \sum_{k=1}^{\infty}\lambda(A_k).$$

Here we also used Exercise 1.40 (monotonicity of λ) and $\tilde{A}_k \subset A_k$.

Taking the infimum over all covers $(A_k)_{k\in\mathbb{N}}$ as above we obtain that

$$\lambda(A) \leq \mu(A),$$

as claimed.

(3) Let $A \in \mathcal{A}$ and $B \subset X$.

Fix $\varepsilon > 0$ arbitrary. By the definition of μ, there exist $(B_k)_{k=1}^{\infty} \subset \mathcal{A}$ such that $B \subset \bigcup_{k=1}^{\infty} B_k$ and

$$\sum_{k=1}^{\infty}\mu(B_k) \geq \mu(B) \geq \sum_{k=1}^{\infty}\mu(B_k)-\varepsilon.$$

Since $B_k \in \mathcal{A}$ we have $\mu(B_k) = \lambda(B_k)$. Since $A, B_k \in \mathcal{A}$ and $\mathcal{A}$ is an algebra we have $B \cap A_k$ and $B_k \backslash A \in \mathcal{A}$. These are disjoint sets, and since λ is a pre-measure we have

$$\mu(B_k) = \lambda(B_k) = \lambda(B_k \cap A) + \lambda(B_k \backslash A).$$

So we have

$$\mu(B) \geq \sum_{k=1}^{\infty}\lambda(B_k \cap A) + \sum_{k=1}^{\infty}\lambda(B_k\backslash A) - \varepsilon.$$

Now observe that $B \cap A \subset \bigcup_{k\in\mathbb{N}} B_k \cap A$ and $B\backslash A \subset \bigcup_{k=1}^{\infty} B_k\backslash A$ (and both coverings belong to $\mathcal{A}$). By the definition of μ we thus find

$$\sum_{k=1}^{\infty} \lambda(B_k \cap A) \geq \mu(B \cap A)$$

and

$$\sum_{k=1}^{\infty} \lambda(B_k \setminus A) \geq \mu(B\backslash A)$$

Together we arrive at

$$\mu(B) \geq \mu(B \cap A) + \mu(B\backslash A).$$

That is, A is μ-measurable.

□

Definition 1.43 (Lebesgue measure)**.** The Lebesgue measure $\mathcal{L}^n$ in $\mathbb{R}^n$ is the Caratheodory-Hahn extension of λ in Example 1.41. Compare this with (1.2).

If the pre-measure is *additionally σ-finite*, then the Carathéodory-Hahn extension μ is essentially unique (on the sets we care about: the measurable sets).

Theorem 1.44 (Uniqueness)**.** *Let $\lambda : \mathcal{A} \to [0, \infty]$ as in Theorem 1.42 be additionally σ-finite and denote by μ the Carathéodory-Hahn-extension.*

Whenever $\tilde{\mu} : 2^X \to [0, \infty]$ is another measure such that

$$\tilde{\mu}(A) = \lambda(A) \quad \textit{for all } A \in \mathcal{A},$$

then indeed

$$\tilde{\mu}(A) = \mu(A) \quad \textit{for all } \mu\textit{-measurable } A$$

Proof. (1) We have $\tilde{\mu}(A) \leq \mu(A)$ for all $A \subset X$. Indeed, let $A \subset \bigcup_{k\in\mathbb{N}} A_k$ for $A_k \in \mathcal{A}$. Then by σ-subadditivity of $\tilde{\mu}$,

$$\tilde{\mu}(A) \leq \sum_{k=1}^{\infty} \tilde{\mu}(A_k) = \sum_{k=1}^{\infty} \lambda(A_k).$$

Taking the infimum over all such covers $(A_k)_{k=1}^{\infty} \subset \mathcal{A}$ of A we obtain

$$\tilde{\mu}(A) \leq \mu(A) \quad \forall A \subset X \tag{1.8}$$

(2) Let Σ be the σ-algebra of μ-measurable sets.
We now show:
$\tilde{\mu}(A) = \mu(A)$ for all $A \subset \Sigma$ such that there is $S \in \mathcal{A}$ with $\lambda(S) < \infty$, and $S \supset A$.
So fix such A and S.
Then

$$\tilde{\mu}(S\backslash A) \leq \tilde{\mu}(S) \overset{S\in\mathcal{A}}{=} \lambda(S) < \infty.$$

Consequently, since $A \in \Sigma$ and $S \in \mathcal{A}$,

$$\tilde{\mu}(A) + \tilde{\mu}(S\backslash A) \overset{(1.8)}{\leq} \mu(A) + \mu(S\backslash A) \overset{A\in\Sigma}{=} \mu(S) \overset{A\in\mathcal{A}}{=} \tilde{\mu}(S) \leq \tilde{\mu}(A) + \tilde{\mu}(S\backslash A).$$

So we have equality everywhere, which leads to

$$\tilde{\mu}(A) + \tilde{\mu}(S\backslash A) = \mu(A) + \mu(S\backslash A)$$

With (1) we conclude

$$\tilde{\mu}(A) + \tilde{\mu}(S\backslash A) \geq \mu(A) + \tilde{\mu}(S\backslash A)$$

and thus $\tilde{\mu}(A) \geq \mu(A)$ (here we use that $\tilde{\mu}(S\backslash A) < \infty$). Again with (1) we have shown that

$$\mu(A) = \tilde{\mu}(A) \quad \forall A \in \Sigma: \quad \text{s.t.} \, \exists \, S \in \mathcal{A}: \; A \subset S, \; \lambda(S) < \infty. \tag{1.9}$$

(3) $\tilde{\mu}(A) = \mu(A)$ for *all* $A \subset \Sigma$

In comparison to (2) we need to remove the restriction $A \subset S$ for some $\lambda(S) < \infty$.

We write $X = \bigcup_{k=1}^{\infty} S_k$ with $(S_k)_{k\in\mathbb{N}}$ pairwise disjoint, $S_k \in \mathcal{A}$, $\lambda(S_k) < \infty$ for all $k \in \mathbb{N}$.

Set $A_k := A \cap S_k$, which are pairwise disjoint sets that all belong to Σ. From (1.9) we have for any $m \in \mathbb{N}$

$$\tilde{\mu}(\bigcup_{k=1}^{m} A_k) = \mu(\bigcup_{k=1}^{m} A_k).$$

Thus, by monotonicity of $\tilde{\mu}$, and since each $A_k \in \Sigma$, we have

$$\tilde{\mu}(A) \geq \limsup_{m\to\infty} \tilde{\mu}(\bigcup_{k=1}^{m} A_k) = \limsup_{m\to\infty} \mu(\bigcup_{k=1}^{m} A_k) \overset{A_k\in\Sigma}{=} \sum_{k=1}^{\infty} \mu(A_k).$$

From Theorem 1.33 (1) we obtain

$$\sum_{k=1}^{\infty} \mu(A_k) = \mu(\bigcup_{k=1}^{\infty} A_k).$$

Thus we have shown for any $A \in \Sigma$,

$$\tilde{\mu}(A) \geq \mu(A),$$

and we have equality in view of (1.8).

□

Remark 1.45. Observe that in Theorem 1.44 it is *not* said that any μ-measurable A was also $\tilde{\mu}$-measurable

Example 1.46. In general μ and $\tilde{\mu}$ in Theorem 1.44 might be different for non-measurable sets.

Let $X = [0, 1]$, $\mathcal{A} = \{\emptyset, X\}$, and set

$$\lambda(\emptyset) := 0, \quad \lambda([0, 1]) := 1.$$

Then $\lambda : \mathcal{A} \to \mathbb{R}$ is a pre-measure.

One can check that the Caratheodory-Hahn extension of λ is given by

$$\mu(A) := \begin{cases} 0 & A = \emptyset \\ 1 & A \neq \emptyset. \end{cases}$$

If on the other hand we consider $\tilde{\mu}$ the Lebesgue measure on $[0,1]$, i.e. the Caratheodory-Hahn extension of $\tilde{\lambda} : \mathcal{A} \to [0,\infty]$ where $\tilde{A}$ are the figures from Example 1.41 and $\tilde{\lambda}$ is the volume as defined. Then $\tilde{\mu}$ coincides with μ on $\mathcal{A}$ but not in $\tilde{A}$, because, e.g.

$$\frac{1}{2} = \lambda([0,1/2]) = \tilde{\mu}([0,1/2]) \neq \mu([0,1/2]) = 1.$$

1.4. Classes of Measures.

Definition 1.47. Let X be a metric space and μ be a measure on X. μ is called a *metric measure* if

$$\mu(E \cup F) = \mu(E) + \mu(F)$$

whenever $\operatorname{dist}(E,F) > 0$

Exercise 1.48. *The Hausdorff measure $\mathcal{H}^s$ on a metric space X is metric measures for any $s \geq 0$.*

Exercise 1.49. *The Lebesgue measure on $\mathbb{R}^n$ is a metric measure.*

Definition 1.50. Let X be a metric space and μ be a measure on X.

- Let $\mathcal{B} \subset 2^X$ be the smallest σ-algebra that contains all open sets of $\mathbb{R}^n$, that is[4]
  $$\mathcal{B} := \bigcap \left\{ \mathcal{A} \subset 2^X : \quad \mathcal{A} \text{ is } \sigma\text{-algebra, all open sets belong to } \mathcal{A} \right\}.$$
  (Cf. Exercise 1.31). Any set $A \in \mathcal{B}$ is called a *Borel set* and $\mathcal{B}$ is called the *Borel σ-algebra*.
- A measure μ on X for which (at least) all Borel-sets are μ-measurable is called a *Borel measure*.

Exercise 1.51. *If $f : X \to Y$ is homeomorphism then $f(A)$ is Borel if and only if A is Borel*

Theorem 1.52. *If μ is a metric measure on a metric space X then μ is a Borel measure, i.e. all open sets are μ-measurable. In particular Lebesgue and Hausdorff measure are Borel measures.*

[4] it is an easy exercise to show that this indeed defines a σ-algebra. Observe in particular that 2^X is a σ-algebra which contains all open sets so the right-hand side is not an intersection of empty sets.

Proof. It suffice to show that all open sets in X are μ-measurable, since then the set of measurable sets (which is a σ-algebra) must contain the Borel sets $\mathcal{B}$.

Let $G \subset X$ be open and $A \subset X$ arbitrary. We need to show

$$\mu(A) \geq \mu(A \cap G) + \mu(A \backslash G).$$

If $\mu(A) = \infty$ this is obvious, so from now on assume $\mu(A) < \infty$.

For $k \in \mathbb{N}$ define

$$G_k := \left\{ x \in G : \quad \operatorname{dist}(x, X \backslash G) > \frac{1}{k} \right\}$$

Then $\operatorname{dist}(G_k, X \backslash G) \geq \frac{1}{k} > 0$.

We have by monotonicity,

$$\mu(A \cap G) + \mu(A \backslash G) \leq \mu(A \cap G_k) + \mu(A \backslash G) + \mu(A \cap (G \backslash G_k)).$$

Since μ is a metric measure and $\operatorname{dist}(A \cap G_k, A \backslash G) > 0$ we conclude

$$\mu(A \cap G) + \mu(A \backslash G) \leq \mu(A) + \mu(A \cap (G \backslash G_k)).$$

So the statement is proven once we show

$$\lim_{k \to \infty} \mu(A \cap (G \backslash G_k)) = 0. \tag{1.10}$$

To see (1.10) let

$$D_k := G_{k+1} \backslash G_k = \left\{ x \in G : \quad \operatorname{dist}(x, X \backslash G) \in (\frac{1}{k+1}, \frac{1}{k}] \right\}.$$

We then have (here we use that G is open)

$$G \backslash G_k = \bigcup_{i=k}^{\infty} D_i \quad \text{so } \mu(A \cap (G \setminus G_k)) \leq \sum_{i=k}^{\infty} \mu(A \cap D_i). \tag{1.11}$$

and whenever $i + 2 \leq j$ we have

$$\operatorname{dist}(D_i, D_j) \geq \frac{1}{i+1} - \frac{1}{j} > 0.$$

Since μ is a metric measure we can sum up even and odd D_i's i.e.

$$\sum_{i=1}^{k} \mu(A \cap D_{2i-1}) = \mu(A \cap \bigcup_{i=1}^{k} D_{2i-1}) \leq \mu(A)$$

and

$$\sum_{i=1}^{k} \mu(A \cap D_{2i}) = \mu(A \cap \bigcup_{i=1}^{k} D_{2i}) \leq \mu(A).$$

In particular we have

$$\sum_{i=1}^{\infty} \mu(A \cap D_i) \leq 2\mu(A) < \infty.$$

From (1.11),

$$\mu(A \cap (G \backslash G_k)) \leq \sum_{i=k}^{\infty} \mu(A \cap D_i)$$

and since the series on the right-hand side converges

$$\mu(A \cap (G \backslash G_k)) \xrightarrow{k \to \infty} 0.$$

That is (1.10) is established and we can conclude. □

Theorem 1.53. *Let X be a metric space and μ any Borel measure. Suppose that X is a union of countably many open sets of finite measure. Then for all Borel sets $E \subset X$*

$$\mu(E) = \inf_{U \supset E; U \text{ open}} \mu(U) = \sup_{C \subset E; C \text{ closed}} \mu(C).$$

The first part of Theorem 1.53 is a direct consequence of

Proposition 1.54. *Let X be a metric space and μ any Borel measure. Suppose that X is a union of countably many open sets of finite measure. If we define $\tilde{\mu} : 2^X \to [0, \infty]$ as*

$$\tilde{\mu}(E) := \inf_{U \supset E; U \text{ open}} \mu(U) \quad E \subset X \tag{1.12}$$

then $\tilde{\mu}$ is a metric measure and we have

$$\tilde{\mu}(E) = \mu(E) \quad \forall \text{Borel sets } E \subset X.$$

The Hausdorff measure $\mathcal{H}^s$ (which is metric and thus Borel) shows that we cannot skip the assumption that X finite union of countably many open sets of finite measure.

Indeed if $s < n$ for any nonempty open set $\mathcal{H}^s(U) = \infty$, so (1.12) is certainly false.

Proof of Proposition 1.54. It is easy to see that $\tilde{\mu}$ is a measure (exercise).

We will show, it is a *metric* measure: let $E, F \subset X$ such that $\delta := \operatorname{dist}(E, F) > 0$. Set

$$V_E := \bigcup_{x \in E} B(x, \frac{1}{3}\delta),$$

and

$$V_F := \bigcup_{x \in F} B(x, \frac{1}{3}\delta).$$

V_E and V_F are then disjoint and open. Fix $\varepsilon > 0$, let $U \supset E \cup F$ be open such that

$$\mu(U) \leq \tilde{\mu}(E \cup F) + \varepsilon.$$

Since $V_E \cap U$ and $V_F \cap U$ are open they are μ-measurable and since they are moreover disjoint

$$\mu((V_E \cap U) \cup (V_F \cap U)) = \mu(V_E \cap U) + \mu(V_F \cap U).$$

Since $E = E \cap V_E \subset U \cap V_E$ and $F = F \cap V_F \subset U \cap V_F$ we have

$$\tilde{\mu}(E) + \tilde{\mu}(F) \leq \mu(V_E \cap U) + \mu(V_F \cap U) = \mu((V_E \cap U) \cup (V_F \cap U)) \leq \mu(U) \leq \tilde{\mu}(E \cup F) + \varepsilon.$$

Taking $\varepsilon \to 0$ we obtain that $\tilde{\mu}(E) + \tilde{\mu}(F) \leq \tilde{\mu}(E \cup F)$ which shows that $\tilde{\mu}$ is metric.

Consequently, in view of Theorem 1.52, $\tilde{\mu}$ is Borel.

It remains to show $\tilde{\mu}(E) = \mu(E)$ for all Borel sets. Clearly,

$$\tilde{\mu}(E) \geq \mu(E) \quad \forall E \subset X.$$

and

$$\tilde{\mu}(E) = \mu(E) \quad \forall E \subset X \text{ open}.$$

By assumption we can write $X = \bigcup_{k=1}^{\infty} X_k$ with X_k open and $\mu(X_k) < \infty$. We also may assume that $X_k \subset X_{k+1}$.

Then we have for all set $E \subset X$

$$\mu(X_n \backslash E) \leq \tilde{\mu}(X_n \backslash E)$$

and

$$\mu(X_n \cap E) \leq \tilde{\mu}(X_n \backslash E).$$

Now if E is Borel, we have

$$\mu(X_n \backslash E) = \tilde{\mu}(X_n \backslash E) \quad \text{and} \quad \mu(X_n \cap E) = \tilde{\mu}(X_n \backslash E). \tag{1.13}$$

Indeed, if not, either $\mu(X_n \backslash E) < \tilde{\mu}(X_n \backslash E)$ or $\mu(X_n \cap E) < \tilde{\mu}(X_n \backslash E)$, which leads to

$$\mu(X_n) = \mu(E \cap X_n) + \mu(X_n \backslash E_n) < \tilde{\mu}(E \cap X_n) + \tilde{\mu}(X_n \backslash E_n) = \tilde{\mu}(X_n) = \mu(X_n)$$

the second to last equation is the $\tilde{\mu}$-measurability of E (since it is a Borel set and $\tilde{\mu}$ is a Borel measure), the last equation uses that X_n is open. The above estimate is impossible (since $\mu(X_n), \tilde{\mu}(X_n) < \infty$), so (1.13) is established.

So in particular we have for any Borel set E, $\mu(X_n \cap E) = \tilde{\mu}(X_n \cap E)$. In view of Theorem 1.33 (recall: both μ and $\tilde{\mu}$ are Borel!)

$$\tilde{\mu}(E) = \lim_{n \to \infty} \tilde{\mu}(X_n \cap E) = \lim_{n \to \infty} \mu(X_n \cap E) = \mu(E).$$

That is, we have shown for any Borel set E,

$$\mu(E) = \tilde{\mu}(E) := \inf_{U \supset E; U \text{ open}} \mu(U). \tag{1.14}$$

□

Proof of Theorem 1.53 last part. Having from Proposition 1.54 (1.14), it remains to prove that for Borel sets E

$$\mu(E) = \sup_{C \subset E; C \text{ closed}} \mu(C).$$

We apply (1.14) to $X_n \backslash E$ and find an open set U_n such that

$$X_n \backslash E \subset U_n$$

and

$$\mu(U_n \backslash (X_n \backslash E)) = \mu(U_n) - \mu(X_n \backslash E) < \frac{\varepsilon}{2^n}.$$

The set $U := \bigcup_{n=1}^{\infty} U_n$ is open and $C = X\backslash U \subset E$ is closed. Now it suffices to observe that

$$E\backslash C = E \cap \bigcup_{n=1}^{\infty} U_n \subset \bigcup_{n=1}^{\infty} G_n\backslash(U_n\backslash E)$$

and hence $\mu(E\backslash C) = \mu(E) - \mu(C) < \varepsilon$. Thus, $\mu(C) \leq \mu(E) \leq \mu(C) + \varepsilon$. Taking $\varepsilon \to 0$ the proof is complete. $\square$

Corollary 1.55. *Let X be a metric space and μ, ν Borel measure. Suppose that X is a union of countably many open sets of finite measure. If ν and μ coincide for open sets, namely if*

$$\nu(U) = \mu(U) \quad \forall \text{open set } U \subset X,$$

then

$$\nu(E) = \mu(E) \quad \forall \text{Borel sets } E.$$

Proof. Twice applying Theorem 1.52, we find that for any Borel set E

$$\mu(E) = \inf_{U \supset E; U \text{ open}} \mu(U) = \inf_{U \supset E; U \text{ open}} \nu(U) = \nu(E).$$

$\square$

On $\mathbb{R}^n$ we can simplify Corollary 1.55

Theorem 1.56 (Borel measures on $\mathbb{R}^n$ that coincide on rectangles)**.** *Let μ and ν be two finite Borel measures on $\mathbb{R}^n$ such that*

$$\mu(R) = \nu(R)$$

for all closed rectangles R of the form

$$R := \{x \in \mathbb{R}^n : \quad a_i \leq x_i \leq b_i, \quad i = 1, \ldots, n\}$$

where $-\infty \leq a_i \leq b_i \leq \infty$, $i = 1, \ldots, n$.

Then

$$\mu(B) = \nu(B)$$

for all Borel sets $B \subset \mathbb{R}^n$

For the proof of Theorem 1.56 we need a essentially combinatorial observation, called the *π-λ Theorem*. π and λ-systems are families of sets which are invariant under less operations than the σ-Algebras.

Definition 1.57.
(1) A nonempty family $\mathcal{P} \subset 2^X$ is called a *π-system* if it is closed under (finitely many) intersections, i.e.

$$A, B \in \mathcal{P} \quad \text{implies} \quad A \cap B \in \mathcal{P}.$$

(2) A family of subsets $\mathcal{L} \subset 2^X$ is called a *λ-system* if
- $X \in \mathcal{L}$
- $A, B \in \mathcal{L}$ *and $B \subset A$* implies $A \setminus B \in \mathcal{L}$

- if $A_k \in \mathcal{L}$ and $A_k \subset A_{k+1}$ *for* $k = 1, \ldots,$ then
$$\bigcup_{k=1}^{\infty} A_k \in \mathcal{L}.$$

Exercise 1.58. *Show that if $\mathcal{P}$ is a λ-system and a π-system, then it is a σ-Algebra.*

Clearly any σ-Algebra is also a π-system and a λ-system.

Theorem 1.59 (π-λ Theorem)**.** *If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system with*
$$\mathcal{P} \subset \mathcal{L}$$
then[5]
$$\sigma(\mathcal{P}) \subset \mathcal{L}$$

Proof. Define $\mathcal{S}$ to tbe the intersection of all λ-systems $\mathcal{L}'$ containing $\mathcal{P}$
$$\mathcal{S} := \bigcap_{\mathcal{L}' \supset \mathcal{P}} \mathcal{L}',$$
We first claim that $\underline{\mathcal{S} \text{ is a } \pi\text{-system}}$.

Indeed let $A, B \in \mathcal{S}$. We must show $A \cap B \in \mathcal{S}$. Define
$$\mathcal{A} := \{C \subset X : \quad A \cap C \in \mathcal{S}\}.$$
Since $\mathcal{S}$ is a λ-system, it follows that $\mathcal{A}$ is a λ-system. Therefore $\mathcal{S} \subset \mathcal{A}$. But then since $B \in \mathcal{S}$ we see that $A \cap B \in \mathcal{S}$.

Next we claim that $\underline{\mathcal{S} \text{ is a } \sigma\text{-algebra}}$

Indeed, we only need to show that $\mathcal{S}$ is a λ-system (cf. Exercise 1.58).

Since $X \in \mathcal{S}$, $\emptyset = X \setminus X \in \mathcal{S}$. Also $A \in \mathcal{S}$ implies $X \setminus A \in \mathcal{S}$. So $\mathcal{S}$ is closed under complements and under finite intersections, and thus under finite unions.

If $A_1, A_2 \in \mathcal{S}$ then $B_n := \bigcup_{k=1}^n A_k \in \mathcal{S}$. Since $\mathcal{S}$ is a λ-system we conclude that $\bigcup_{k=1}^\infty A_k = \bigcup_{n=1}^\infty B_n \in \mathcal{S}$. Thus $\mathcal{S}$ is a σ-algebra.

Since $\mathcal{S} \supset \mathcal{P}$ is a σ-algebra it follows that
$$\sigma(\mathcal{P}) \subset \mathcal{S} \subset \mathcal{L}.$$
□

Proof of Theorem 1.56. Let
$$\mathcal{P} := \{A \subset \mathbb{R}^n : \quad \text{for some } n \in \mathbb{N},\ A = \bigcap_{k=1}^n R_k \text{ where } (R_k)_{k=1}^n \text{ are a rectangles}\},$$

[5] Recall that $\sigma(\mathcal{P})$ is the σ-Algebra generated by $\mathcal{P}$

which a π-system. We also set

$$\mathcal{L} := \{B \subset \mathbb{R}^n : \quad B \text{ Borel:} \quad \mu(B) = \nu(B)\}.$$

While it is not so clear that $\mathcal{L}$ is a σ-Algebra, one can check it is a λ-system. Also $\mathcal{P} \subset \mathcal{L}$ by assumption (observe that each R_k is μ and σ-measurable).

By Theorem 1.59 we have $\sigma(\mathcal{P}) \subset \mathcal{L}$. Since $\sigma(\mathcal{P})$ contains the Borel sets (any open set can be written as countable union of rectangles, see Lemma 1.75 below), any Borel set B satisfies $B \in \mathcal{L}$ and thus $\mu(B) = \nu(B)$, and we can conclude. $\square$

Definition 1.60 (Borel Regular measure)**.** A Borel measure μ is *Borel regular*, if for any $A \subset \mathbb{R}^n$ there exists some Borel set $B \supset A$ such that $\mu(A) = \mu(B)$.

Corollary 1.61. *Let X be a metric space and μ any Borel measure. Suppose that X is a union of countably many open sets of finite measure. If we define $\tilde{\mu} : 2^X \to [0,\infty]$ as*

$$\tilde{\mu}(E) := \inf_{U \supset E; U \text{ open}} \mu(U) \quad E \subset X$$

then $\tilde{\mu}$ is a metric, Borel-regular measure that coincides with μ on Borel sets.

Proof. In view of Proposition 1.54 we only need to show that $\tilde{\mu}$ is Borel-regular.

Indeed, let $E \subset X$ with $\mu(E) < \infty$ (otherwise $\mu(E) = \mu(X) = \infty$) be arbitary. Let $(U_k)_{k=1}^\infty$ be open sets so that $U_k \supset E$ and

$$\mu(U_k) - \frac{1}{k} \leq \tilde{\mu}(E).$$

Set $U := \bigcap_{k=1}^\infty U_k$. Then we have $E \subset U$ and thus $\tilde{\mu}(E) \leq \tilde{\mu}(U)$. On the other hand we have

$$\tilde{\mu}(U) = \tilde{\mu}(\bigcap_{k=1}^\infty U_k) \leq \mu(U_k) \leq \tilde{\mu}(E) + \frac{1}{k}.$$

This holds for any $k \in \mathbb{N}$ so letting $k \to \infty$ we find

$$\tilde{\mu}(U) \leq \tilde{\mu}(E).$$

So $\tilde{\mu}(U) = \tilde{\mu}(E)$, and since as an intersection of countably many open sets U is a Borel-set, we can conclude. $\square$

Example 1.62. The Lebesgue measure $\mathcal{L}^n$ (as defined in Definition 1.43) is Borel regular.

Proof. In view of Corollary 1.61 it suffices to show that

$$\mathcal{L}^n(A) = \inf \{\mathcal{L}^n(G) : \quad G \subset \mathbb{R}^n \text{ open and } G \supset A\}.$$

$\leq$ is obvious by monotonicity. For $\geq$ assume $\mathcal{L}^n(A) < \infty$, let $\varepsilon > 0$ and take blocks $(Q_i)_{i=1}^\infty$ such that

$$\sum_{i=1}^\infty \text{vol}(Q_i) \leq \mathcal{L}^n(A) + \varepsilon.$$

To each block Q_i we can choose an open block $Q_i^o \supset Q_i$ such that
$$\mathrm{vol}(Q_i^o) \leq \mathrm{vol}(Q_i) + 2^{-i}\varepsilon.$$
Set $G := \bigcup_{i=1}^\infty (Q_i)^o \supset A$. Then
$$\begin{aligned}\mathcal{L}^n(G) \leq& \sum_{i=1}^\infty \mathrm{vol}(Q_i^o) \leq \sum_{i=1}^\infty \mathrm{vol}(Q_i) + \sum_{i=1}^\infty 2^{-i}\varepsilon\\ \leq& \mathcal{L}^n(A) + 2\varepsilon.\end{aligned}$$
That is we have shown
$$\mathcal{L}^n(A) + 2\varepsilon \geq \inf\left\{\mathcal{L}^n(G) : \quad G \subset \mathbb{R}^n \text{ open and } G \supset A\right\}.$$
which holds for any $\varepsilon > 0$, letting $\varepsilon \to 0$ we conclude. □

Example 1.63. Let X be a metric space. For any $s \geq 0$, $\mathcal{H}^s(X)$ is Borel-regular.

Proof. In this case we cannot use Corollary 1.61, since we cannot assume that we can cover X with countably many finite measure sets.

If $s = 0$ the claim is easy. Take any $E \subset X$. If $\mathcal{H}^0(E) = \infty$, then $\mathcal{H}^0(E) = \mathcal{H}^0(X) = \infty$. If $\mathcal{H}^0(E) < \infty$ then E contains finitely many points, thus E is closed and thus a Borel set.

Now let $s > 0$ and $E \subset X$. If $\mathcal{H}^s(E) = \infty$ then again $\mathcal{H}^s(E) = \mathcal{H}^s(X) = \infty$ and we conclude. So assume $\mathcal{H}^s(E) < \infty$. In particular $\mathcal{H}^s_\delta(E) < \infty$ for all $\delta > 0$.

So for each ℓ there exists a covering of E by open balls $(B(r_{k;\ell}))_{k=1}^\infty$ with radius $r_{k;\ell} \leq \frac{1}{\ell}$ such that
$$\mathcal{H}^s_{\frac{1}{\ell}}\Big(\bigcup_{k=1}^\infty B(r_{k;\ell})\Big) \leq \mathcal{H}^s_{\frac{1}{\ell}}(E) + \frac{1}{\ell} \leq \mathcal{H}^s(E) + \frac{1}{\ell}. \tag{1.15}$$
In the last step we used again $\mathcal{H}^s(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A)$.

Then $G := \bigcap_{\ell=1}^\infty \bigcup_{k=1}^\infty B(r_{k;\ell})$ is a Borel set, and we have
$$\mathcal{H}^s(G) = \lim_{\ell\to\infty} \mathcal{H}^s_{\frac{2}{\ell}}(G) \leq \limsup_{\ell\to\infty} \mathcal{H}^s_{\frac{2}{\ell}}\Big(\bigcup_{k=1}^\infty B(r_{k;\ell})\Big) \overset{(1.15)}{\leq} \mathcal{H}^s(E) + 0$$
Since on the other hand $G \supset E$ we have
$$\mathcal{H}^s(G) = \mathcal{H}^s(E).$$
We can conclude. □

Definition 1.64. Let X be a metric space and $\mu : 2^X \to [0,\infty]$ a Borel regular measure. μ is called a *Radon measure* if $\mu(K) < \infty$ for all compact sets.

Example 1.65.
- $\mathcal{L}^n$ is a Radon measure on $\mathbb{R}^n$
- $\mathcal{H}^s$ is not a Radon measure in $\mathbb{R}^n$ whenever $s < n$.

Example 1.66. Let X is a locally compact and separable metric space[6] and μ is a Borel measure such that $\mu(K) < \infty$ for all K compact.

Then X is the union of countably many open sets of finite measure and thus μ coincides with a Radon measure $\tilde{\mu}$ on all Borel sets; cf. Corollary 1.61.

Exercise 1.67. *Let X be a metric space and μ any Radon measure. Suppose that X is a union of countably many open sets of finite μ-measure. Let $A \subset X$ be μ-measurable. Show that $\mu \llcorner A$ is a Radon measure.*

Theorem 1.68. *Let X be a metric space and μ any Borel-regular measure. Suppose that X is a union of countably many open sets of finite μ-measure.*

(1) If μ is Borel-regular, then for any $A \subset X$ we have

$$\mu(A) = \inf_{A \subset G,\ G \text{ open}} \mu(G)$$

(2) If $X = \mathbb{R}^n$ and μ is a Radon measure, for any μ-measurable $A \subset X$ we have

$$\mu(A) = \sup_{F \subset A,\quad F \text{ compact}} \mu(F).$$

Proof. (1) Let $A \subset X$. By monotonicity we always have

$$\mu(A) \leq \inf_{A \subset G,\ G \text{ open}} \mu(G).$$

For the other inequality, since μ is by assumption Borel-regular we find a Borel set $E \supset A$ such that $\mu(A) = \mu(E)$. By Theorem 1.53 we have

$$\mu(A) = \mu(E) = \inf_{E \subset G,\ G \text{ open}} \mu(G) \geq \inf_{A \subset G,\ G \text{ open}} \mu(G)$$

(2) Follows from part (1), see Exercise 1.69.

□

Exercise 1.69. *Show Theorem 1.68 (2).*

Use the argument as in the proof of Theorem 1.53 last part and Theorem 1.68(1). Observe we need measurability of A because we need to use that $\mu(U \backslash (X_n \backslash A)) = \mu(U) - \mu(X_n \backslash A)$, which is the measurability of $X_n \backslash A$.

Theorem 1.70. *Let X be a metric space μ a Radon measure on X, and assume that X is a union of countably many open sets of finite μ-measure. Then the following are equivalent for $A \subset X$.*

(1) A is μ-measurable
(2) For every $\varepsilon > 0$ there is an open set $G \subset X$ such that $A \subset G$ and $\mu(G \backslash A) < \varepsilon$

[6] Separable means, a countable set is dense. Locally compact means that ever point has a neighborhood whose closure is compact. $\mathbb{R}^n$ is locally compact and separable, but also $\mathbb{R}^n \setminus 0$ is locally compact and separable

(3) There is a G_δ-set*, namely a set* $H = \bigcap_{i=1}^\infty G_i$ *for some open sets* $G_i \subset \mathbb{R}^n$*, such that* $A \subset H$ *and* $\mu(H \backslash A) = 0$
(4) For every $\varepsilon > 0$ *there is a closed set* F *such that* $F \subset A$ *and* $\mu(A \backslash F) < \varepsilon$
(5) There is a F_σ-set*, namely a set* $M = \bigcup_{i=1}^\infty F_i$ *for some closed set* $F_i \subset \mathbb{R}^n$*, such that* $M \subset A$ *and* $\mu(A \backslash M) = 0$
(6) For every $\varepsilon > 0$ *there is an open set* G *and a closed set* F *such that* $F \subset A \subset G$ *and* $\mu(G \backslash F) < \varepsilon$.
(7) $A = B \cup N$ *where* B *is a Borel set and* N *is a* μ*-zero-set.*

Proof. By assumption we have $X = \bigcup_{k=1}^\infty X_k$ with X_k open and $\mu(X_k) < \infty$.

(1) $\Rightarrow$ (2): Let $\varepsilon > 0$. Set $A_k := A \cap X_k$ then $\mu(A_k) < \infty$ and by Theorem 1.68 there exists for each k an open set $G_k \supset A_k$ such that

$$\mu(G_k) \leq \mu(A_k) + \frac{\varepsilon}{2^k}.$$

Since A_k is measurable and $G_k \supset A_k$ we have

$$\mu(G_k \backslash A_k) = \mu(G_k) - \mu(A_k) \leq \frac{\varepsilon}{2^k}.$$

Hence $A \subset G := \bigcup_{k=1}^\infty G_k$ and we have

$$\mu(G \setminus A) \leq \sum_{k=1}^\infty \mu(G_k \setminus A) \leq \sum_{k=1}^\infty \mu(G_k \setminus A_k) \leq \sum_{k=1}^\infty \frac{\varepsilon}{2^k} = \varepsilon.$$

Letting $\varepsilon \to 0$ we conclude. (2) $\Rightarrow$ (3). We set $H := \bigcap_{i=1}^\infty U_i$ where U_i are open sets such that $A \subset U_i$ and $\mu(U_i \backslash A) < \frac{1}{i}$. Then $A \subset H$ and

$$\mu(H \setminus A) \leq \mu(U_i \setminus A) \leq \frac{1}{i} \xrightarrow{i \to \infty} 0.$$

(3) $\Rightarrow$ (1) Let $A \subset H$ for some G_δ-set H with $\mu(H \setminus A) = 0$. Then for any $B \subset X$ we have by monotonicity

$$\mu(B \cap A) \leq \mu(B \cap H)$$

and since $B \cap H \subset (B \cap A) \cup (H \backslash A)$,

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

thus we actually have

$$\mu(B \cap H) = \mu(B \cap A).$$

Similarly,

$$\mu(B \setminus H) \leq \mu(B \setminus A)$$

and since $B \setminus A \subset (B \setminus H) \cup (H \setminus A)$,

$$\mu(B \setminus A) \leq \mu(B \setminus H) + \mu(H \setminus A) = \mu(B \setminus H).$$

Thus

$$\mu(B \setminus H) = \mu(B \setminus A).$$

Now since any G_δ-set is a Borel set and μ is in particular a Borel measure we have that H is μ-measurable. Consequently

$$\mu(B) = \mu(B \cap H) + \mu(B \setminus H) = \mu(B \cap A) + \mu(B \setminus A).$$

Thus A is μ-measurable.

$\underline{(1) \Leftrightarrow (4) \Leftrightarrow (5)}$ this equivalence follows from the equivalence of the conditions (1), (2), and (3) applied to $X \backslash A$.

$\underline{(1) \Rightarrow (6)}$ If A is μ-measurable, the existence of the sets F and G follows from the conditions (2) and (4).

$\underline{(6) \Rightarrow (3)}$ Take closed and open sets F_i, G_i such that $F_i \subset A \subset G_i$, $\mu(G_i \backslash F_i) < \frac{1}{i}$. Then the set $H := \bigcap_{i=1}^\infty G_i$ is G_δ and $\mu(H \backslash A) = 0$.

$\underline{(7) \Rightarrow (1)}$ If $A = B \cup N$ where B is a Borel set and N is a μ-zero-set, then clearly A is it is the union of two μ-measurable sets.

$\underline{(5) \Rightarrow (7)}$ We have $A = M \cup A \backslash M$, where M is a F_σ-set, in particular a Borel set. and $A \setminus M$ is a μ-zero set. $\square$

From Theorem 1.56 we obtain the following uniqueness result

Theorem 1.71 (Radon measures on $\mathbb{R}^n$ that coincide on rectangles)**.** *Let μ and ν be two Radon measures on $\mathbb{R}^n$ such that*

$$\mu(R) = \nu(R)$$

for all closed rectangles R of the form

$$R := \{x \in \mathbb{R}^n : \quad a_i \leq x_i \leq b_i, \quad i = 1, \ldots, n\}$$

where $-\infty \leq a_i \leq b_i \leq \infty$, $i = 1, \ldots, n$.

Then

$$\mu(A) = \nu(A) \quad \forall A \subset \mathbb{R}^n.$$

Proof. By Theorem 1.56 we have

$$\mu(K) = \nu(K)$$

for all compact set K (observe that then the measure is w.l.o.g. finite). By Theorem 1.68(2) we obtain that

$$\mu(A) = \nu(A)$$

for all Borel sets (which are both μ- and ν-measurable).

If $A \subset \mathbb{R}^n$ is any set, then there exists two Borel sets $B_\mu \supset A$ and $B_\nu \supset A$ such that

$$\mu(A) = \mu(B_\mu)$$

and

$$\nu(A) = \nu(B_\nu).$$

Observe that we have

$$\mu(A) \leq \mu(B_\nu \cap B_\mu) = \nu(B_\nu \cap B_\mu) \leq \nu(B_\nu) = \nu(A).$$

Similarly we find $\nu(A) \leq \mu(A)$. That is $\mu(A) = \nu(A)$ and we can conclude. □

Proposition 1.72. *Let μ be a Radon measure on $\mathbb{R}^n$. Then for each $x \in \mathbb{R}^n$ there are at most countably many $r > 0$ such that*

$$\mu(\partial B(x,r)) > 0.$$

Exercise 1.73. *Construct a Radon measure μ on $\mathbb{R}^n$ such that for countably many $r > 0$,*

$$\mu(\partial B(x,r)) > 0.$$

Proof of Proposition 1.72. For simplicity let us assume that $x = 0$.

Now consider

$$f(r) := \mu(\overline{B(r)}).$$

Clearly $f : (0,\infty) \to [0,\infty)$ is increasing, so f has at most countably many points of discontinuity.

Moreover we observe that

$$\mu(B(r)) = \lim_{\tilde{r} \to r^-} f(\tilde{r}).$$

Indeed, this follows since for any increasing sequence $r_1 < r_2 < \ldots < r$ with $\lim r_i = r$, so by Theorem 1.33(2)

$$\mu(B(r)) = \mu(\bigcup_{i=1}^\infty \overline{B(r_i)}) = \lim_{i\to\infty} \mu(\overline{B(r_i)}) = \lim_{i\to\infty} f(r_i).$$

Next we claim that whenever r is a point of continuity for f then $\mu(\partial B(r)) = 0$. Indeed, since $\overline{B(r)}$ and $\partial B(r)$ are compact and $B(r)$ open (thus all are measurable),

$$0 \leq \mu(\partial B(r)) = \mu(\overline{B(r)}) - \mu(B(r)) = f(r) - \lim_{\tilde{r}\to r^-} f(\tilde{r}).$$

So if r is a point of continuity, then $\lim_{\tilde{r}\to r^-} f(\tilde{r}) = f(r)$ and thus

$$\mu(\partial B(r)) = 0.$$

□

Similarly we have

Exercise 1.74. *Let μ be a Radon measure on $\mathbb{R}^n$ and $g \in C^0(\mathbb{R}^n, \mathbb{R})$ such that $g \equiv 0$ outside a compact set K. Then there exist at most countable $r \in \mathbb{R}$ such that*

$$\mu(g^{-1}(r)) > 0.$$

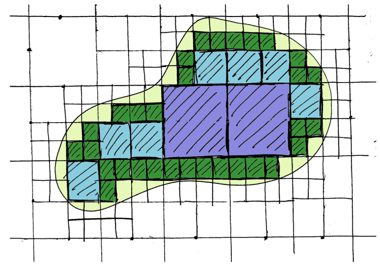


FIGURE 1.5. An open set can be represented by a countable union of dyadic cubes.

1.5. More on the Lebesgue measure.

Lemma 1.75 (Dyadic decomposition of open sets). *Any open set $\Omega \subset \mathbb{R}^n$ can be written as*

$$\Omega = \bigcup_{i=1}^{\infty} Q_i$$

where Q_i are closed cubes with pairwise disjoint interior and each cube has sidelength 2^{-k} for some $k \in \mathbb{N}$.

Proof. See Figure 1.5 for a picture proof. We consider *dyadic cubes* of sidelength $2^{-\ell}$, $\ell \in \mathbb{N} \cup \{0\}$ which cover $\mathbb{R}^n$:

$$\mathbb{R}^n = \bigcup_{a \in 2^{\ell}\mathbb{Z}^n} a + [0, 2^{-\ell}]^n.$$

Let

$$\mathcal{Q}_\ell := \left\{ Q = a + [0, 2^{-\ell}]^n : \quad a \in 2^{\ell}\mathbb{Z}^n \text{ and } Q \subset \Omega \setminus \left(\bigcup_{\tilde{Q} \in \mathcal{Q}_{\ell-1}} \tilde{Q} \right)^o \right\}.$$

Each $\mathcal{Q}_\ell$ is a countable family of cubes, so $\mathcal{Q} := \bigcup_{\ell=0}^{\infty} \mathcal{Q}_\ell$ is a countable family of cubes.

By construction we have

$$\bigcup_{Q \in \mathcal{Q}} Q \subset \Omega.$$

Observe also that by construction two cubes in $\mathcal{Q}$ intersect only along their boundary.

Now assume by contradiction that $\bigcup_{Q \in \mathcal{Q}} Q \neq \Omega$, i.e.

$$\Omega \setminus \bigcup_{Q \in \mathcal{Q}} Q \supset \{x_0\}.$$

Observe that any point $x \in \bigcup_{Q \in \mathcal{Q}} Q$ belongs to at most 2^n many cubes $Q \in \mathcal{Q}$.

We conclude that $\bigcup_{Q \in \mathcal{Q}} Q$ is a closed set, even though it may be a countable union of closed cubes. Indeed assume $(x_i)_{i \in \mathbb{N}} \subset \bigcup_{Q \in \mathcal{Q}} Q$ converges to some $\bar{x}$. Then all but finitely many x_i belong to the same cube $\tilde{Q}$ which is closed, so $\bar{x} \in \tilde{Q}$.

By the argument from above, $\Omega \setminus \bigcup_{Q\in\mathcal{Q}} Q$ is open thus there exists a small ball $B(x_0,\rho) \subset \Omega \setminus \bigcup_{Q\in\mathcal{Q}} Q \supset \{x_0\}$. Let k_0 such that $2^{-k_0} < \frac{\rho}{1000}$. Then we can find a tiny cube $\bar{Q} = a + [0, 2^{-k_0}] \subset B(x_0,\rho)$ (each component of x_0 lies between some $[2^{-k_0(\gamma+1)}, 2^{-k_0(\gamma)}]$). Contradiction because that cube would belong to $\tilde{Q}$. $\square$

Exercise 1.76. *Show that any open set $U \subset \mathbb{R}^n$ can be written as a countable union of open cubes.*

The Lebesgue measure is essentially built from blocks, the dyadic decomposition shows that the Lebesgue measure is then the only natural volume notion on $\mathbb{R}^n$: any other translation invariant volume notion is the same.

Theorem 1.77 (Uniqueness of the Lebesgue measure)**.** *If μ is a Borel-measure on $\mathbb{R}^n$ such that*

- *$\mu(a+E) = \mu(E)$ for all $a \in \mathbb{R}^n$ and all Borel-sets $E \subset \mathbb{R}^n$ (translation invariance)*
- *$\mu([0,1]^n) = 1$*

then $\mu(E) = \mathcal{L}^n(E)$ for all Borel-sets E.

If μ is moreover Borel-regular (and thus a Radon measure) then $\mu(E) = \mathcal{L}^n(E)$ for all $E \subset \mathbb{R}^n$.

Proof. Consider an open cube $Q_k = (0, 2^{-k})^n$ where k is a positive integer. There are 2^{kn} pairwise disjoint cubes contained in the unit cube $[0,1]^n$, each being a translation of Q_k. Since the measure μ is invariant under translation we have

$$\mu(Q_k) \leq 2^{-kn}.$$

Now we can use a similar argument to cover ∂Q by translations small cubes in such a way that the sum of μ-measures of the cubes covering ∂Q goes to zero as $k \to \infty$. Thus $\mu(\partial Q) = 0$ for any cube Q.

Now $[0,1]^n$ can be covered by 2^{kn} cubes which are translations of $\overline{Q_k}$. So,

$$1 = \mu([0,1]^n) \leq 2^{kn}\mu(\overline{Q_k}).$$

Thus, $\mu(\overline{Q_k}) \geq 2^{-kn}$ and $\mu(Q_k) \leq 2^{-kn}$. Since $\mu(\partial Q_k) = 0$ we find that $\mu(\overline{Q_k}) = \mu(Q_k) = 2^{-kn}$ for all k.

In conclusion: any cube Q of sidelength 2^{-k} satisfies

$$\mu(Q) = \mathcal{L}^n(Q).$$

By dyadic decomposition, Lemma 1.75, any open set E can be written as $E = \bigcup_{i=1}^{\infty} Q_i$ for cubes which are translations of Q_k for some k, and with pairwise disjoint interior. Thus

$$\mu(E) = \sum_i \mu(Q_i) = \sum_i \mathcal{L}^n(Q_i) = \mathcal{L}^n(E).$$

In view of Theorem 1.53 applied to μ and the Lebesgue measure we conclude that $\mu(E) = \mathcal{L}^n(E)$ for all Borel sets E.

If μ is Borel regular then the additional claim follows from Theorem 1.71. □

Since the n-Hausdorf measure $\mathcal{H}^n$ a translation invariant, Exercise 1.16, Borel measure, we obtain that it essentially coincides with the Lebesgue measure.

Corollary 1.78. *Let $n \in \mathbb{N}$. Then n-th Hausdorff measure and Lebesgue measure coincide in $\mathbb{R}^n$ up to a multiplicative factor C_n for all Borel sets. That is*

$$\mathcal{H}^n(E) = \mathcal{L}^n(E) \quad \forall E \subset \mathbb{R}^n.$$

Proof. From Example 1.63, Exercise 1.16, Theorem 1.77 we have $\mathcal{H}^n(E) = C\mathcal{L}^n(E)$ for some constant C. The fact that $C = 1$ one can check in by testing for $E = B(0,1)$. □

Proposition 1.79. *A set $E \subset \mathbb{R}^n$ has Lebesgue measure zero if and only if for every $\varepsilon > 0$ there is a family of balls $(B(x_i, r_i))_{i=1}^\infty$ such that*

$$E \subset \bigcup_{i=1}^\infty B(x_i, r_i)$$

and

$$\sum_{i=1}^\infty r_i^n < \varepsilon.$$

Proof. $\Leftarrow$. Assume

$$E \subset \bigcup_{i=1}^\infty B(x_i, r_i) \tag{1.16}$$

and

$$\sum_{i=1}^\infty r_i^n < \varepsilon. \tag{1.17}$$

Obserserve that if $Q(x_i, 2r_i)$ denotes the cube with sidelength $2r_i$ centered at x_i then $B(x_i, r_i) \subset Q(x_i, 2r_i)$ and thus by the definition of the Lebesgue measure, Definition 1.43,

$$\mathcal{L}^n(B(x_i, r_i)) \leq \mathrm{vol}(Q(x_i, 2r_i)) = (4r_i)^n.$$

Consequently, by monotonicity

$$\mathcal{L}^n(E) \leq \sum_{i=1}^\infty 4^n r_i^n \leq 4^n \varepsilon.$$

Thus if for every $\varepsilon > 0$ there is a family of balls $(B(x_i, r_i))_{i=1}^\infty$ such that (1.16) and (1.16) holds, we have

$$\mathcal{L}^n(E) = 0.$$

That is E is a $\mathcal{L}^n$-zero set.

$\underline{\Rightarrow}$: Assume $\mathcal{L}^n(E) = 0$.

By definition of the Lebesgue measure, Definition 1.43, for any $\varepsilon > 0$ there exist $(A_k)_{k=1}^\infty$ figures which each can decomposed into blocks of pairwise disjoint interior $A_k = \bigcup_{i=1}^{N_k} Q_{k;i}$, such that

$$E \subset \bigcup_{k=1}^\infty A_k \equiv \bigcup_{k=1}^\infty Q_{k;i}$$

and

$$\sum_{k=1}^\infty \mathrm{vol}(A_k) \equiv \sum_{k=1}^\infty \sum_{i=1}^{N_k} \mathrm{vol}(Q_{k;i}) < \varepsilon.$$

Denote by $L_{k;i}$ the sidelength of the cube $Q_{k;i}$, and by $x_{k;i}$ the center of the cube $Q_{k;i}$. Then $Q_{k;i} \subset B(x_{k;i}, L_{k;i})$ and

$$\mathrm{vol}(Q_{k;i}) = c_n (L_{k;i})^n$$

for some dimensional constant c_n. Thus,

$$E \subset \bigcup_{k}^\infty \bigcup_{i=1}^{N_k} B(x_{k;i}, L_{k;i})$$

and

$$\sum_{k=1}^\infty \sum_{i=1}^\infty (L_{k;i})^n = \frac{1}{c_n} \sum_{k=1}^\infty \sum_{i=1}^\infty \mathrm{vol}(Q_{k;i}) < \frac{\varepsilon}{c_n}.$$

□

Exercise 1.80. *Use Proposition 1.79 to show that*

(1) *if $f : \mathbb{R}^n \to \mathbb{R}^n$ is uniformly continuous, then for any $A \subset \mathbb{R}^n$ with $\mathcal{L}^n(A) = 0$ we also have $\mathcal{L}^n(f(A)) = 0$. This property of f is called the Lusin property.*
(2) *Whenever $\Sigma \subset \mathbb{R}^n$ with $\mathcal{L}^n(\Sigma) = 0$ then $\mathbb{R}^n \setminus \Sigma$ is dense in $\mathbb{R}^n$, i.e. $\overline{\mathbb{R}^n \setminus \Sigma} = \mathbb{R}^n$.*

So we know the Lebesgue measure is translation invariant (and that its pretty much the only measure that does that). But we have not verified that this means that it is rotation invariant. Namely what is the $\mathcal{L}^n$ measure of a cube Q rotated? The following result shows (in particular) that rotations do not change the measure.

Theorem 1.81. *Let $Lx := Ax + b$ be an affine linear map $\mathbb{R}^n \to \mathbb{R}^n$ with a non-degenerate matrix $A \in \mathbb{R}^{n\times n}$, i.e. $\det(A) \neq 0$, and $b \in \mathbb{R}^n$. Then*

(1) *$E \subset \mathbb{R}^n$ is a Borel set if and only if $L(E) \subset \mathbb{R}^n$ is a Borel set.*
(2) *E is $\mathcal{L}^n$-measurable if and only if $L(E)$ is $\mathcal{L}^n$-measurable*
(3) *We have $\mathcal{L}^n(L(\Omega)) = |\det(L)|(\Omega)$ for all $\Omega \subset \mathbb{R}^n$.*

Proof. (1) L is a homeomorphism, so $E \subset \mathbb{R}^n$ is a Borel set if and only if $L(E) \subset \mathbb{R}^n$ is a Borel set, see Exercise 1.51.

(2) In view of Theorem 1.70(7), a set A is $\mathcal{L}^n$-measurable if and only if $A = B \cup N$ where B is a Borel set and N is a zero set. Since L and L^{-1} are both Lipschitz maps they map zero sets into zero sets, Exercise 1.13 and Corollary 1.78. Since L and L^{-1} also preserve the Borel property of set we get the claim.

(3) By Theorem 1.68 applied to the Lebesgue measure $\mathcal{L}^n$ it suffices to show

$$\mathcal{L}^n(L(\Omega)) = \det(L)(\Omega) \text{ for all open } \Omega \subset \mathbb{R}^n$$

Let

$$\mu(A) := \mathcal{L}^n(L(A)) \quad A \in \mathbb{R}^n.$$

Then μ is a Borel measure, and it is still translation invariant. So we apply Theorem 1.77 to $\tilde{\mu} := \frac{1}{a}\mu$ where $a := \mu([0,1]^n)$ and have that

$$\mathcal{L}^n(L(\Omega)) = \mathcal{L}^n(L([0,1]^n))\,\mathcal{L}^n(\Omega) \quad \text{for all open sets } \Omega.$$

So it remains to show that

$$\mathcal{L}^n(A([0,1]^n)) = |\det(A)| \quad \forall A \in GL(n).$$

So consider $f : GL(n) \to (0,\infty)$, $f(A) := \mathcal{L}^n(A([0,1]^n))$. We observe two properties: first,

$$f(AB) = f(A)f(B) \quad \forall A, B \in GL(n). \tag{1.18}$$

Indeed, we have

$$f(AB) = \mathcal{L}^n(AB([0,1]^n)) = f(A)\mathcal{L}^n(B([0,1]^n)) = f(A)f(B)\mathcal{L}^n(([0,1]^n)).$$

Secondly, for the identity matrix $I \in \mathbb{R}^{n\times n}$

$$f(\lambda I) = \lambda^n, \quad \forall \lambda > 0. \tag{1.19}$$

This is immediate from the definition of the pre-measure of the Lebesgue measure.

We conclude with the following Lemma, Lemma 1.82

□

Lemma 1.82. *Let $f : GL(n) \to (0,\infty)$ satisfy (1.18) and (1.19). Then*

$$f(A) = |\det(A)|.$$

Proof. Let $A_i(s)$ be the diagonal matrix with $-s$ on the ith place and s on all other places on the diagonal. Since $A_i(s)^2 = s^2 I$ we have $f(A_i(s))^2 = f(A_i(s)^2) = s^{2n}$ and hence $f(A_i(s)) = |s^n| = |\det A_i(s)|$. For $k \neq \ell$ let $B_{k\ell}(s) = (a_{ij})_{ij}$ be the matrix such that $a_{k\ell} = s$, $a_{ii} = 1$, $i = 1, 2, \ldots, n$ and all other entries equal zero. Multiplication by the matrix $B_{kl}(s)$ from the right (left) is equivalent to adding kth column (ℓth row) multiplied by s to ℓth column (kth row).

It is well known from linear algebra (and easy to prove) that applying such operations to any nonsingular matrix A it can be transformed to a matrix of the form tI or $A_n(t)$. Since multiplication by $B_{k\ell}(s)$ does not change determinant, $t = |\det A|^{1/n}$. It remains to prove that $f(B_{kl}(s)) = 1$. Since $B_{kl}(-s) = A_k(1)\,B_{kl}(s)\,A_k(1)$ we have that $f(B_{kl}(-s)) =$

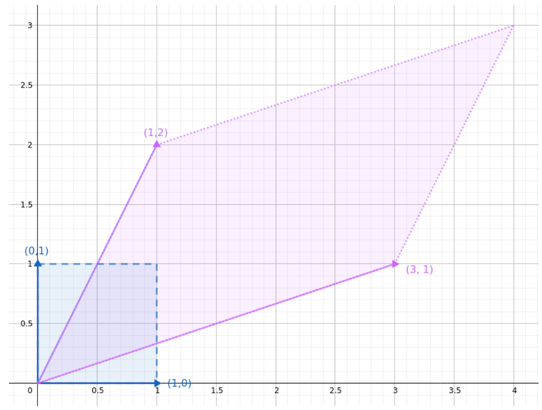


FIGURE 1.6. Let $a_{11} = 3$, $a_{12} = 1$, $a_{21} = 1$ and $a_{22} = 2$ then $[0,1]^2$ (blue) is transformed into $A[0,1]^2$ (purple). The purple area is $\det(A) = 5$.

$f(B_{kl}(s))$. On the other hand $B_{k\ell}(s)B_{k\ell}(-s) = I$ and hence $f(B_{kl}(s))^2 = 1$, so $f(B_{kl}(s)) = 1$. The proof of Lemma 1.82 and hence the proof of Theorem 1.81 is complete. $\square$

Example 1.83. Let $A \in \mathbb{R}^{n\times n}$, and let $I = [0,1]^n$. Set

$$AI := \{Ax, x \in I\}.$$

Then the volume (in the elementary geometrical sense) of AI is

$$|AI| = |\det(A)||I|.$$

Exercise 1.84. *Let $Lx := Ax + b$ be an affine linear map $\mathbb{R}^n \to \mathbb{R}^n$ for any matrix $A \in \mathbb{R}^{n\times n}$ and any $b \in \mathbb{R}^n$. Show the following:*

(1) If E is $\mathcal{L}^n$-measurable then $L(E)$ is $\mathcal{L}^n$-measurable
(2) We have $\mathcal{L}^n(L(\Omega)) = |\det(L)|(\Omega)$ for all $\Omega \subset \mathbb{R}^n$.

Hint: If $\det(A) \neq 0$ this follows from Theorem 1.81. If $\det(A) = 0$, what is $L(E)$ or $L(\Omega)$? (compare to the Hausdorff measure $\mathcal{H}^s$ for $s = \operatorname{rank} A$)

Corollary 1.85. *The Lebesgue measure $\mathcal{L}^n$ is translation and rotationally invariant on $\mathbb{R}^n$. Namely if $A \subset \mathbb{R}^n$ then $\mathcal{L}^n(A) = \mathcal{L}^n(\Phi(A))$ where for some $x_0 \in \mathbb{R}^n$ and $R \in O(n)$ we have*

$$\Phi(x) := x_0 + Rx$$

then $\mathcal{L}^n(\Phi(A)) = \mathcal{L}^n(A)$.

1.6. **Nonmeasurable sets.** Corollary 1.85 combined with the Banach-Tarski paradoxon, (1.1) and Figure 1.1, implies the existence of nonmeasurable sets (w.r.t. $\mathcal{L}^n$) in $\mathbb{R}^n$: Any set A and B such that for some pairwise disjoint and *measurable* C_i they are represented by

$$A = \bigcup_{i=1}^N C_i, \quad \text{and} \quad B = \bigcup_{i=1}^N (x_i + O_i C_i).$$

(where $O_i \in O(n)$ is a rotation and x_i a point, and $(x_i + O_iC_i)_{i=1}^N$ are pairwise disjoint) would satisfy

$$\mathcal{L}^n(A) = \sum_{i=1}^N \mathcal{L}^n(C_i) = \sum_{i=1}^N \mathcal{L}^n(x_i + O_iC_i) = \mathcal{L}^n(B),$$

so if $A = [0,1]^n$ and $B = [0,2]^n$ this would mean the impossible $\mathcal{L}^n(A) = \mathcal{L}^n(B)$ – so one of the C_i must be nonmeasurable.

We will not prove the Banach-Tarski-paradox, but instead show the $\mathbb{R}^1$-version.Observe below that the existence of nonmeasurable sets requires the the axiom of choice (and suitable invariances of the underlying measure).

Theorem 1.86 (Vitali)**.** *Let $\mu : 2^{\mathbb{R}} \to \mathbb{R}$ be translation invariant*[7]*, i.e.*

$$\mu(x + A) = \mu(A) \quad \forall x \in \mathbb{R}, A \subset \mathbb{R}.$$

If moreover $\mu([0,1]) \in (0,\infty)$ then there exists a Vitali-set $A \subset [0,1]$ that is not μ-measurable.

Proof. Construction (Vitali) Fix $\xi \in \mathbb{R}\backslash\mathbb{Q}$, and set

$$G_\xi := \{k + \ell\xi; \quad k, \ell \in \mathbb{Z}\}$$

We use G_ξ to define an equivalence relation $\sim$ on $\mathbb{R}$.

$$x \sim y :\Leftrightarrow x - y \in G_\xi.$$

For $x \in \mathbb{R}$ denote by $[x]$ the set

$$[x] := x + G_\xi = \{y \in \mathbb{R} : \quad y = x + k + \ell\xi \quad k, \ell \in \mathbb{Z}\}.$$

Let $A \subset \mathbb{R}$ be a set such that for each class $[x]$ there exists exactly one element $y \in A \cap [x]$. The set A exists by the *axiom of choice*: if we set

$$X := \{[x] \subset \mathbb{R} : \quad x \in \mathbb{R}\}$$

then the axiom of choice says there exists a choice function $f : X \to \mathbb{R}$ such that $f([x]) \in [x]$ for all $[x] \in X$. Then $A := f(X)$.

Without loss of generality, $A \subset [0,1]$. Indeed if we can adapt the choice function f above such that

$$\tilde{f}([x]) := f([x]) - k,$$

where $k \in \mathbb{Z}$ is chosen such that $f([x]) \in [k, k+1)$.

The Vitali-set A is not μ-measurable

Since $\xi \not\in \mathbb{Q}$ the set $G_\xi \cap [-1,1]$ contains infinitely many points. Indeed: if $k + \ell\xi = k' + \ell'\xi$ then $k = k'$ and $\ell = \ell'$ (otherwise $\xi = \frac{k-k'}{\ell'-\ell} \in \mathbb{Q}$). Thus we can find infinitely many different $k + \ell\xi \in [-1,1]$ (choosing $k = -\lfloor \ell\xi \rfloor$ or similar)

[7] yes, if μ is moreover Borel then it is pretty much the Lebesgue measure, Theorem 1.77, but the main point of this theorem is: invariances mean non-measurable sets

Also, G_ξ is clearly countable we can write $G_\xi \cap [-1,1] = \bigcup_{k=1}^\infty g_k$ where $(g_k)_{k=1}^\infty$ are pairwise different points. Set $A_k := g_k + A \subset [-1,2]$, $k \in \mathbb{N}$.

We claim

$$A_k \cap A_\ell = \emptyset \quad k \neq \ell. \tag{1.20}$$

Indeed, assume that $x \in A_k \cap A_\ell$ then

$$x - g_k, x - g_\ell \in A.$$

Observe that $x - g_k \sim x - g_\ell$, that is $[x] = [x - g_k] = [x - g_\ell]$. But by definition of A there is exactly one element of $[x - g_k] = [x - g_\ell]$ in A, which implies that $x - g_k = x - g_\ell$, that is $g_k = g_\ell$, that is $k = \ell$. This establishes (1.20).

Next we claim that

$$[0,1] \subset \bigcup_{k=1}^\infty A_k. \tag{1.21}$$

Indeed, let $y \in [0,1]$. By the construction of A there must be exactly one $x \in A \cap [y]$. In particular $y - x \in G_\xi$. Since $x, y \in [0,1]$ we have $y - x \in [-1,1] \cap G_\xi$, and thus there must be some $g_k = y - x$. Consequently

$$y = x + g_k \subset A + g_k = A_k.$$

This establishes (1.21).

Now assume A is μ-measurable, i.e.

$$\mu(B) = \mu(A \cap B) + \mu(B \setminus A) \quad \forall B \subset \mathbb{R}.$$

Then by translation invariance, so is A_k, indeed for any $B \subset \mathbb{R}$,

$$\begin{aligned}
\mu(B) =& \mu(B - g_k) = \mu(A \cap (B - g_k)) + \mu((B - g_k) \setminus A) \\
=& \mu(g_k + A \cap (B - g_k)) + \mu(g_k + (B - g_k) \setminus A) \\
=& \mu((A + g_k) \cap B) + \mu(B \setminus (A + g_k)) \\
=& \mu(A_k \cap B) + \mu(B \setminus A_k).
\end{aligned}$$

Then we have by (1.20) and σ-additivity of disjoint measurable sets, Theorem 1.33,

$$\sum_{k=1}^\infty \mu(A_k) = \mu(\bigcup_{k=1}^\infty A_k).$$

Since $[0,1] \subset \bigcup A_k \subset [-1,2]$ we have again by translation invariance

$$\mu([0,1]) \leq \sum_{k=1}^\infty \mu(A_k) \leq \mu([0,3]) \leq \mu([0,1]) + \mu([1,2]) + \mu([2,3]) = 3\mu([0,1]).$$

On the other hand, again by translation invariance

$$\mu(A_k) = \mu(A).$$

That is if A was measurable we'd have

$$\mu([0,1]) \leq \sum_{k=1}^{\infty} \mu(A) \leq 3\mu([0,1]).$$

This is only possible if $\mu([0,1]) = \infty$ or $\mu([0,1]) = 0$, both are ruled out by assumption. □

So in general even for very 'reasonable' measures μ (in the sense that they measure indeed something like area, volume, length etc.) we cannot really hope to avoid the existence of non-measurable sets. The way to deal with that fact is to shun non-measurable sets, and only work with measurable sets.

This strategy will propagate throughout this course: we only care about functions that stay (in a reasonable way) within the category of measurable sets.

2. Measurable functions

Our goal is integration, and for this purpose it is convenient with functions that can be infinite (think of $\frac{1}{|x|^2}$). We will use the notation

$$\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}.$$

Definition 2.1. Let (X, Σ, μ) be a measure space. We say that $f : X \to \overline{\mathbb{R}}$ is a *measurable function*, if f^{-1} maps open sets of $\mathbb{R}$ into μ-measurable sets (where we consider $\{+\infty\}$, $\{-\infty\}$ as open sets). That is

$$f^{-1}(U) \in \Sigma \quad \forall \text{open sets } U \subset \mathbb{R} \text{ and } U = \{+\infty\} \text{ and } U = \{-\infty\}.$$

More generally if Ω is μ-measurable then $f : \Omega \to Y$ is a measurable function if $f^{-1}(U)$ is μ-measurable for any open set $U \subset Y$ (and $U = \{+\infty\}$ and $U = \{-\infty\}$)

For consistency we observe

Lemma 2.2. *Let (X, Σ, μ) be a measure space $\Omega \subset X$ μ-measurable and $f : \Omega \to \overline{\mathbb{R}}$. Recall the notation of the measure $\mu\llcorner\Omega$,*

$$\mu\llcorner\Omega(A) := \mu(\Omega \cap A).$$

Denote the $\mu\llcorner\Omega$-measurable sets by Σ_Ω. Then

$$\Sigma_\Omega := \{A \subset \Omega : \quad A \ \mu\text{-measurable}\} = \{\Omega \cap B : \quad B \ \mu\text{-measurable}\}$$

And in particular the following are equivalent

(1) f is μ-measurable with respect to (X, Σ, μ)
(2) f is $\mu\llcorner\Omega$-measurable with respect to $(\Omega, \Sigma, \mu\llcorner\Omega)$.

Exercise 2.3. *Let X be some set, μ a measure on X. Assume $f : X \to \overline{\mathbb{R}}$.*

Show that

$$\mathcal{G} := \left\{ B \subset \mathbb{R} : f^{-1}(B) : \quad \text{is } \mu\text{-measurable} \right\} \subset 2^{\mathbb{R}}$$

is a σ-Algebra in $\mathbb{R}$.

Lemma 2.4. *Let (X, Σ, μ) be a measure space $\Omega \subset X$ μ-measurable and $f : \Omega \to \overline{\mathbb{R}}$.*

The following are equivalent

(1) *$f^{-1}(U)$ is μ-measurable for any open set $U \subset \mathbb{R}$*
(2) *$f^{-1}(B)$ is μ-measurable for any Borel set $B \subset \mathbb{R}$*
(3) *$f^{-1}((-\infty, a])$, $f^{-1}((-\infty, a))$, $f^{-1}((-\infty, a])$ $f^{-1}([a, \infty))$ are μ-measurable for any $a \in \mathbb{R}$.*

Proof. (1) $\Rightarrow$ (2) Observe that

$$\mathcal{G} := \left\{ B \subset \mathbb{R} : f^{-1}(B) : \quad \text{is } \mu\text{-measurable} \right\}$$

is a σ-Algebra (Exercise 2.3) which in view of (1) contains the open sets. So $\mathcal{G}$ contains the Borel sets.

(2) $\Rightarrow$ (1) and (2) $\Rightarrow$ (3) are obvious since any open set and any interval is a Borel set.

(3) $\Leftrightarrow$ (2) follows the same way as above, since the Borel σ-algebra are generated by the infinite intervals. □

We can (usually we don't want to) extend a bit the notion of measurability to a *topological space* as target.

Definition 2.5 (Topological space)**.** A *topological space* is a set X and a collection $\tau \subset 2^X$ of "open sets" which satisfies the following axioms

- $\emptyset \in \tau$ and $X \in \tau$.
- Let I be a (finite or infinite) index set and let $A_i \in \tau$ for all $i \in I$. Then $\bigcup_{i \in I} A_i \in \tau$.
- Let $N \in \mathbb{N}$ and $A_i \in \tau$ for $i \in \{1, \ldots, N\}$ then

$$\bigcap_{i=1}^{N} A_i \in \tau.$$

τ is called a *topology*.

For two topological spaces (X, τ_X) and (Y, τ_Y) a map $f : X \to Y$ is *continuous* if f^{-1} maps open sets to open sets, i.e. $f^{-1}(U) \in \tau_X$ for any $U \in \tau_Y$.

If (X, d) is a metric space then the collection of open sets (w.r.t. the metric) form a topology.

Continuity between two metric spaces (w.r.t. topology) is the same as continuity w.r.t metric structure.

Remark 2.6. When talking about measurability of a map $f : \Omega \to \overline{\mathbb{R}}$ we actually consider $\overline{\mathbb{R}}$ as a *topological space*, where the open sets $A \subset \overline{\mathbb{R}}$ are

- the open sets in $\mathbb{R}$
- the "open neighborhoods" of $+\infty$ and $-\infty$, i.e. $[-\infty, a)$ and $(a, \infty]$ for $a \in \mathbb{R}$

and unions thereof.

Then $f : \Omega \to \overline{\mathbb{R}}$ is μ-measurable if and only if $f^{-1}(U)$ is μ-measurable whenever $U \subset \overline{\mathbb{R}}$ is open.

In particular we have that $f : \Omega \to \overline{\mathbb{R}}$ is *measurable if and only if $f^{-1}([-\infty, a))$ and $f^{-1}((a, +\infty])$ are measurable for all $a \in \mathbb{R}$.*

This leads to the following definition:

Definition 2.7. Let (X, Σ, μ) be a measure space, $\Omega \subset X$ measurable, and (Y, τ) a topological space. A map $f : \Omega \to Y$ is called measurable if $f^{-1}(U)$ is measurable for any open set $U \subset Y$ (i.e. for any set $U \in \tau$).

Example 2.8. Let (X, d) be a metric space and μ be a Borel-measure. Then any continuous function $f : X \to \mathbb{R}$ is μ-measurable.

Proof. $f^{-1}(+\infty) = f^{-1}(-\infty) = \emptyset$ which is clearly μ-measurable. Moreover since f is continuous $f^{-1}(U)$ is open whenever U is open, so $f^{-1}(U)$ is μ-measurable since μ is Borel. $\square$

Exercise 2.9. *Let (X, Σ, μ) be a measure space, Y and Z topological spaces. Assume $f : X \to Y$ is a measurable function and $g : Y \to Z$ is continuous. Then $g \circ f : X \to Z$ is measurable.*

Example 2.10. Let (X, Σ, μ) be a measure space and Y a topological space. If the functions $u_1, u_2, \dots, u_n : X \to \mathbb{R}$ are measurable and $\Phi : \mathbb{R}^n \to Y$ is continuous, then the function

$$h(x) = \Phi(u_1(x), u_2(x), \dots, u_n(x)) : X \to Y$$

is μ-measurable.

Proof. We only discuss the case $n = 2$, and write $u_1 = u$ and $u_2 = v$.

Set $f(x) := (u(x), v(x))$. In view of Exercise 2.9 it suffices to prove that f is measurable. We observe that if $R = (a, b) \times (c, d)$ is an open rectangle, then

$$f^{-1}(R) = u^{-1}((a, b)) \cap u^{-1}((c, d)) \text{is measurable.}$$

Since any open set $U \subset \mathbb{R}^2$ can be written as a countable union of open rectangles (Exercise 1.76))

$$U = \bigcup_{i=1}^{\infty} R_i$$

we have

$$f^{-1}(U) = \bigcup_{i=1}^{\infty} f^{-1}(R_i),$$

so $f^{-1}(U)$ is μ-measurable. □

Exercise 2.11. *(1) If $f = (f^1, f^2, \dots, f^n) : X \to \mathbb{R}^n$ where $f^1, \dots, f^n : X \to \mathbb{R}$ are μ-measurable, then f is μ-measurable (take $\Phi(f^1, \dots, f^n) := (f^1, \dots, f^n)^t$).*

(2) If $f : X \to \mathbb{R}^n$ is μ-measurable, then its components $f^i : X \to \mathbb{R}$ are μ-measurable, $i = 1, \dots, n$. (take $\Phi_i(f) := f^i$).

Also $|f|$ is μ-measurable (take $\Phi(f) := |f|$).

(3) If $f, g : X \to \mathbb{R}$ are μ-measurable then so are $f + g$, $f - g$, $-f$, and $fg : X \to \mathbb{R}$ [8]

Example 2.12. (1) A set $E \subset X$ is μ-measurable if and only if its *characteristic function* χ_E

$$\chi_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise} \end{cases}$$

is μ-measurable. Indeed

$$(\chi_E)^{-1}(U) = \begin{cases} \emptyset & \text{if } 0 \notin U \text{ and } 1 \notin U \\ E & \text{if } 0 \notin U \text{ and } 1 \in U \\ X \backslash E & \text{if } 0 \in U \text{ and } 1 \notin U \\ X & \text{if } 0 \in U \text{ and } 1 \in U \end{cases}$$

(2) There are non-$\mathcal{L}^1$-measurable functions $f : \mathbb{R} \to \mathbb{R}$. Indeed take the Vitali-Set V from Theorem 1.86. Then $f := \chi_V$ is not measurable.

Theorem 2.13. *Let $f_k : \Omega \to \overline{\mathbb{R}}$ μ-measurable, $k \in \mathbb{N}$. Then*

$$F_1(x) := \inf_k f_k(x)$$

$$F_2(x) := \sup_k f_k(x)$$

$$F_3(x) := \liminf_{k\to\infty} f_k(x)$$

$$F_4(x) := \limsup_{k\to\infty} f_k(x)$$

are all μ-measurable.

[8] observe that product is pointwise defined since $f(x), g(x) \neq \pm\infty$, so we have no problem with $0 \cdot \infty$

Proof. Observe that

$$(\inf_{k\in\mathbb{N}} f_k)^{-1}([-\infty,a)) = \bigcup_{k=1}^{\infty} f_k^{-1}([-\infty,a)),$$

$$(\inf_{k\in\mathbb{N}} f_k)^{-1}((a,\infty]) = \bigcap_{k=1}^{\infty} f_k^{-1}((a,\infty]),$$

so F_1 is measurable. Since

$$\sup_k f_k(x) = -\inf_k(-f_k)(x)$$

we see that F_2 is measurable. For F_3 and F_4 observe

$$\liminf_{k\to\infty} f_k(x) = \lim_{\ell\to\infty}\left(\inf_{k\geq\ell} f_k(x)\right) = \sup_{\ell\in\mathbb{N}}\left(\inf_{k\geq\ell} f_k(x)\right),$$

and similarly

$$\limsup_{k\to\infty} f_k = \inf_{\ell\in\mathbb{N}}\left(\sup_{k\geq\ell} f_k\right).$$

□

The following theorem says that any μ-measurable function can be approximated by *simple functions*. Simple functions (sometimes called *step functions*) are functions

$$f(x) = \sum_{i=1}^{n} \lambda_i \chi_{A_i}$$

where $\lambda_i \neq \lambda_j$ for $i \neq j$ and A_i are disjoint measurable sets.

Theorem 2.14. *Let $f : (X,d) \to [0,\infty]$ be a μ-measurable function. Then there are μ-measurable sets $A_k \subset X$, for all $k \in \mathbb{N}$ such that*[9]

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \quad \forall x \in X.$$

Proof. Let

$$A_1 := \{x \in X : \quad f(x) \geq 1\} = f^{-1}([1,\infty]).$$

Since f is μ-measurable, A_1 is μ-measurable. Now define inductively the μ-measurable sets

$$A_k := \left\{x \in X : \quad f(x) \geq \frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j}\chi_{A_j}(x)\right\} = \left(f - \sum_{j=1}^{k-1}\frac{1}{j}\chi_{A_j}\right)^{-1}([\frac{1}{k},\infty]), \quad k = 2,3,\ldots.$$

Now let $x \in X$. We first show

$$f(x) \geq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \quad \forall x \in X. \tag{2.1}$$

[9] observe that the sum of the right hand side either converges absolutely or is infinite

If $\underline{\#\{k \in \mathbb{N}: \; x \in A_k\} = \infty}$ then $f(x) \geq \sum_{j=1}^{k} \frac{1}{j}\chi_{A_j}(x)$ for infinitely many k, and thus

$$f(x) \geq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x).$$

If $\underline{1 \leq \#\{k \in \mathbb{N}: \quad x \in A_k\} < \infty}$ then set $k_0 := \max\{k \in \mathbb{N}: \quad x \in A_k\} \in \mathbb{N}$. Since $x \in A_{k_0}$ and $x \notin A_k$ for $k > k_0$ we have

$$f(x) \geq \frac{1}{k_0} + \sum_{j=1}^{k_0-1} \frac{1}{j}\chi_{A_j}(x) = \frac{1}{k_0} \underbrace{\chi_{A_{k_0}}(x)}_{=1} + \sum_{j=1}^{k_0-1} \frac{1}{j}\chi_{A_j}(x) + \sum_{j=k_0}^{\infty} \frac{1}{j} \underbrace{\chi_{A_j}(x)}_{=0} = \sum_{j=1}^{\infty} \frac{1}{j}\chi_{A_j}(x).$$

If $\underline{\{k \in \mathbb{N}: \quad x \in A_k\} = \emptyset}$ then

$$f(x) \geq 0 = \sum_{k=1}^{\infty} \frac{1}{k} \underbrace{\chi_{A_k}(x)}_{=0}.$$

(2.1) is now established.

We can conclude once we show that also

$$f(x) \leq \sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \quad \forall x \in X. \tag{2.2}$$

Let $x \in X$. If $\underline{f(x) = \infty}$ then $x \in A_k$ for all k, so

$$f(x) = \infty = \sum_{k=1}^{\infty} \frac{1}{k} = \sum_{k=1}^{\infty} \frac{1}{k} \underbrace{\chi_{A_k}(x)}_{=1}.$$

If $\underline{f(x) = 0}$ (2.2) is obvious since the right-hand side is nonnegative.

If $\underline{0 < f(x) < \infty}$, then we may assume that $x \notin A_k$ for infinitely many $k \in \mathbb{N}$. Indeed, otherwise there exists some k_0 such that $x \in A_k$ for all $k \geq k_0$ and thus

$$\sum_{k=1}^{\infty} \frac{1}{k}\chi_{A_k}(x) \geq \sum_{k=k_0}^{\infty} \frac{1}{k} \underbrace{\chi_{A_k}(x)}_{=1} = \sum_{k=k_0}^{\infty} \frac{1}{k} = \infty \geq f(x),$$

and (2.2) is established.

So the only remaining case is that $f(x) \in (0, \infty)$ and $x \notin A_k$ for infinitely many $k \in \mathbb{N}$. Now take an increasing sequence $k_i \to \infty$ with $x \notin A_{k_i}$ for each i. With the definition of A_k we then have

$$f(x) \leq \frac{1}{k_i} + \sum_{j=1}^{k_i-1} \frac{1}{j}\chi_{A_j}(x) \quad \forall i.$$

That is,

$$f(x) \leq \limsup_{k \to \infty} \left(\frac{1}{k} + \sum_{j=1}^{k-1} \frac{1}{j}\chi_{A_j}(x) \right) = \sum_{j=1}^{\infty} \frac{1}{j}\chi_{A_j}(x)$$

(2.2) is now established. □

Another way to approximate by simple functions is

Exercise 2.15. *Let $f : X \to [0, \infty)$ be μ-measurable and bounded*

$$\sup_{x \in X} f(x) < \infty$$

Set

$$s_n(x) := \begin{cases} n & \text{if } f(x) \geq n \\ \frac{k}{2^k} & \text{if } \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n} \leq n. \end{cases}$$

Show that s_n converges uniformly to f.

So we can approximate nonnegative measurable function $f : X \to [0, \infty]$ by *step functions*,

$$f(x) = \lim_{L \to \infty} \sum_{k=1}^{L} \frac{1}{k} \chi_{A_k}(x).$$

- Observe that this approximation is monotone increasing.
- Splitting a μ-measurable function $f : X \to \bar{\mathbb{R}}$ into $f = f_+ - f_-$ and we can approximate any measurable function by step-functions.

Exercise 2.16. *Let X be a metric space. A function $f : X \to \bar{R}$ is* upper semicontinuous

$$\limsup_{x \to x_0} f(x) \leq f(x_0) \quad \forall \text{for all } x_0 \in X$$

A function is lower semicontinuous *if*

$$\liminf_{x \to x_0} f(x) \geq f(x_0) \quad \text{for all } x_0 \in X.$$

(1) Show that upper semicontinuity is equivalent to saying

$$f^{-1}([a, \infty]) \text{is closed} \quad \forall a \in \mathbb{R}.$$

and this in turn is equivalent to

$$f^{-1}([-\infty, a)) \text{is open} \quad \forall a \in \mathbb{R}.$$

(2) Show that any step function $f(x) = \sum_{k=1}^{K} \chi_{A_k}$ is upper semicontinuous if A_k are closed

(3) Show that any step function $f(x) = \sum_{k=1}^{K} \chi_{A_k}$ is lower semicontinuous if A_k are open

(4) Show that if μ is Borel measure then any upper or lower semicontinuous function is μ-measurable.

3. Integration

In integration theory we will follow the notation that $0 \cdot \infty = 0$.

Definition 3.1. Let X be a metric space and μ a measure.

- Let $f : X \to [0, \infty)$ be a (nonnegative!) simple function of the form
$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x)$$
where $\alpha_i \neq \alpha_j$ for $i \neq j$ and A_i are pairwise disjoint μ-measurable sets.
For any μ-measurable set $\Omega \subset X$ we define
$$\int_\Omega f\, d\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i \cap \Omega).$$
(Recall that if $\mu(A) = \infty$ we still, by assumption, set $0 \cdot \mu(A) = 0$. And if $\alpha = \infty$ but $\mu(A_i \cap \Omega) = 0$ then still $\alpha\mu(A_i \cap \Omega) := 0$.)
- It is easy to see that if f is a simple function represented by two different sums
$$f(x) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(x) = \sum_{j=1}^{m} \beta_j \chi_{B_j}(x)$$
then
$$\sum_{i=1}^{n} \alpha_i \mu(A_i \cap \Omega) = \sum_{j=1}^{m} \beta_j \mu(B_i \cap \Omega),$$
so the integral notion is well-defined.
- Clearly if $f(x) \le g(x)$ for all x and both f and g are simple functions, then
$$\int_\Omega f d\mu \le \int_\Omega g d\mu.$$
- Thus, if $f : X \to [0, \infty)$ is a (nonnegative!) simple function as above then
$$\int_\Omega f\, d\mu = \sup_s \int_\Omega s d\mu$$
where the supremum is taken over all simple functions s such that $0 \le s \le f$.
- So we can extend the above notion to all nonnegative measurable functions. If $g : X \to [0, \infty]$ is a μ-measurable function and $\Omega \subset X$ is μ-measurable we define the *Lebesgue integral* of g over Ω by
$$\int_\Omega g d\mu := \sup_s \int_\Omega s d\mu$$
where the supremum is taken over all simple functions s such that $0 \le s \le g$ a.e.

It is worth to note that this defines the Lebesgue integral for all nonnegative measurable functions (it may just be infinite). This is reminiscient of the lower Darboux sum approximation of the Riemann integral, the main difference here being that we use step functions not on blocks but on measurable sets.

We observe the following properties

Exercise 3.2.
- *If* $0 \leq f \leq g$ *then* $\int_E f d\mu \leq \int_E g d\mu$
- *If* $A \subset B$ *and* $f \geq 0$ *then* $\int_A f d\mu \leq \int_B f d\mu$.
- *If* $f(x) = 0$ *for all* $x \in E$ *then* $\int_E f d\mu = 0$, *even if* $\mu(E) = \infty$.
- *If* $\mu(E) = 0$ *then* $\int_E f d\mu = 0$ *even if* $f(x) = \infty$ *for all* $x \in E$.
- *if* $f \geq 0$ *then* $\int_E d\mu = \int_X \chi_E f \, d\mu$.

The following is an easy observation

Exercise 3.3. *Let* g *and* g *be nonnegative,* μ*-measurable simple functions on* X *and* $\lambda \geq 0$ *a constant. Then*

$$\int_X (f+g) d\mu = \int_X f d\mu + \int_X g d\mu$$

and

$$\int_X \lambda f d\mu = \lambda \int_X f d\mu.$$

It is easy to conclude homogeneity

Exercise 3.4 (homogeneity of the integral)**.** *Let* $f : X \to [0, \infty]$ *be* μ*-measurable and* $\lambda \geq 0$*. Show, without using the Lebesgue monotonce convergence theorem, that*

$$\int_X \lambda f d\mu = \lambda \int_X f d\mu.$$

So we have reason to believe the integral is linear, i.e.

$$\int_X \lambda f + g d\mu = \lambda \int_X f d\mu + \int_X g d\mu.$$

We will prove this below, Corollary 3.7, with a technique that can be generalized to one of the most fundamental theorems of integration theory, the *Lebesgue monotone convergence theorem*

Theorem 3.5 (Lebesgue monotone convergence theorem)**.** *Let* $(f_n)_n$ *be a sequence of* μ*-measurable functions on* X *such that*

- $0 \leq f_1 \leq f_2 \leq \ldots \leq \infty$ *for almost every*[10] $x \in X$
- $\lim_{n\to\infty} f_n(x) = f(x) \in [0, \infty]$ *for almost every* $x \in X$.

Then f *is* μ*-measurable and*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Before we can prove Theorem 3.5 we need the following

[10] recall Definition 1.35

Lemma 3.6. *Let f be a nonnegative μ-measurable simple functions on X. Denote by Σ the μ-measurable sets. The map $\tilde{\nu} : \Sigma \to [0, \infty]$ given by*

$$\tilde{\nu}(E) := \int_E f \, d\mu \quad \forall E \ \mu\text{-measurable}$$

is a premeasure, cf. Definition 1.39.

Its Carathéodory-Hahn Extension ν of $\tilde{\nu}$, cf. Theorem 1.42, in symbols

$$\nu := f \llcorner \mu,$$

is called the concatenation of f and μ.

In particular any μ-measurable set E is $f \llcorner \mu$-measurable.

Proof. Since Σ is a σ-Algebra, it is is in particular an algebra.

To confirm that $\tilde{\nu}$ is a premeasure, let $A = \bigcup_{k=1}^{\infty} A_k$, where $(A_k)_{k=1}^{\infty}$ are pairwise disjoint μ-measurable functions.

We need to show

$$\tilde{\nu}(A) = \sum_{k=1}^{\infty} \tilde{\nu}(A_k).$$

Let f be given as

$$f = \sum_{i=1}^{n} \alpha_i \chi_{B_i}.$$

Then

$$\tilde{\nu}(A_k) = \int_{A_k} f d\mu = \sum_{i=1}^{n} \alpha_i \mu(B_i \cap A_k).$$

For each $i \in \{1, \ldots, n\}$, the collection $(A_k \cap B_i)_{k \in \mathbb{N}}$ consists of measurable and pairwise disjoint sets. By Theorem 1.33(1) we find

$$\sum_{k=1}^{\infty} \tilde{\nu}(A_k) = \sum_{k=1}^{\infty} \sum_{i=1}^{n} \alpha_i \mu(B_i \cap A_k) = \sum_{i=1}^{n} \alpha_i \mu(B_i \cap A) = \int_A f d\mu = \tilde{\nu}(A).$$

This proves that $\tilde{\nu}$ is a pre-measure. □

Proof of Theorem 3.5. f is measurable in view of Theorem 2.13. We can replace the "almost every" in the assumption by "every", by changing f_i and f on zero sets (observe that a countable union of zeroset is a zeroset).

By monotonicity of $\int d\mu$ the sequence

$$\alpha_n := \int_X f_n d\mu \in [0, \infty]$$

is monotonically increasing and hence it has a limit $\alpha = \lim_{n\to\infty} \int_X f_n d\mu \in [0,\infty]$. Since $f_n(x) \leq f(x)$ everywhere we have

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu \leq \int_X f d\mu. \tag{3.1}$$

To conclude we need to show "$\geq$" in the above inequality.

Let $0 \leq s \leq f$ be a simple function. For a fixed constant $0 < c < 1$ we define

$$E_n := \{x \in X : \quad f_n(x) \geq c\, s(x)\}, \quad n \in \mathbb{N}.$$

Since f_n is increasing, $E_1 \subset E_2 \subset \ldots$. Since $c < 1$ and $f_n(x) \xrightarrow{n\to\infty} f(x)$ for each x we have that $\bigcup_n E_n = X$.

Since the E_n are μ-measurable sets ($f_n(\cdot) - cs(\cdot)$ is measurable!) we have

$$\int_X f_n d\mu \geq \int_{E_n} f_n d\mu \geq c \int_{E_n} s d\mu = cs\llcorner\mu(E_n). \tag{3.2}$$

Now we have by Lemma 3.6 that $s\llcorner\mu$ is a measure and each E_n is $s\llcorner\mu$-measurable. Since $E_n \subset E_{n+1}$ we find by Theorem 1.33

$$s\llcorner\mu(E_n) \xrightarrow{n\to\infty} s\llcorner\mu(\bigcup_{n=1}^{\infty} E_n) = s\llcorner\mu(X) = \int_X s d\mu.$$

Taking the limit in (3.2) we have

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu \geq c \int_X s d\mu.$$

Recall that this holds whenever s is a simple function with $s \leq f$. Taking the supremum over such s we conclude

$$\alpha \geq c \int_X f d\mu.$$

This holds for any $c \in (0,1)$, taking the limit as $c \to 1$ we find

$$\alpha \geq \int_X f d\mu.$$

Combining this with (3.1) we conclude. $\square$

Corollary 3.7 (Linearity of the integral)**.** *Let $f, g : X \to [0,\infty]$ μ-measurable and $\lambda, \sigma \geq 0$. Then*

$$\int_X \lambda f + \sigma g d\mu = \lambda \int_X f d\mu + \sigma \int_X g d\mu.$$

Proof. We only show

$$\int_X f + g d\mu = \int_X f d\mu + \int_X g d\mu.$$

the general case then follows from this and Exercise 3.4.

Let s_i and t_i be *monotonically increasing* sequence of simple nonnegative functions such that $s_i \xrightarrow{i\to\infty} f$ and $t_i \xrightarrow{i\to\infty} f_2$ (existence follows from Theorem 2.14). Then $s_i + t_i \xrightarrow{i\to\infty} f + g$.

From Lebesgue monotone convergence theorem, Theorem 3.5 and Exercise 3.3

$$\int_X f + g d\mu = \lim_{i\to\infty} \int_X (s_i + t_i) d\mu = \lim_{i\to\infty} \int_X s_i d\mu + \int_X t_i d\mu = \int_X f d\mu + \int_X g d\mu$$

□

Indeed, we have more than Corollary 3.7 we can take out infinite sums (recall that all functions considered are *nonnegative*, so series always absolute converge (or are infinity))

Corollary 3.8. *Let $f_n : X \to [0, \infty]$ be a sequence of μ-measurable functions and set for $x \in X$*

$$f(x) := \sum_{n=1}^{\infty} f_n(x) \in [0, \infty].$$

Then f is measurable and we have

$$\int_X f d\mu = \sum_{i=1}^{\infty} \int_X f_n d\mu.$$

Proof. By induction and Corollary 3.7 we have

$$\sum_{i=1}^{N} \int_X f_n d\mu = \int_X \sum_{i=1}^{N} f_n d\mu$$

Now observe that f_n are nonnegative so $\sum_{i=1}^N f_n$ is a monotonce sequence in N with $\lim_{N\to\infty} \sum_{i=1}^N f_n = \sum_{i=1}^\infty f_n$. So by Lebesgue monotone convergence we conclude that

$$\sum_{i=1}^{\infty} \int_X f_n d\mu = \lim_{N\to\infty} \sum_{i=1}^{N} \int_X f_n d\mu = \lim_{N\to\infty} \int_X \sum_{i=1}^{N} f_n d\mu = \int_X \sum_{i=1}^{\infty} f_n d\mu.$$

□

Then next important consequence is called *Fatou's lemma*, which essentially tells us *lower semi-continuity of the integral under pointwise convergence*

Corollary 3.9 (Fatou's Lemma)**.** *Let $f_n : X \to [0, \infty]$ be a sequence of μ-measurable functions, then*

$$\int_X \left(\liminf_{n\to\infty} f_n\right) d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

Proof. Let $g_n := \inf\{f_n, f_{n+1}, \ldots\}$. Then $g_n \leq f_n$ and hence

$$\int_X g_n d\mu \leq \int_X f_n d\mu$$

Since $0 \leq g_1 \leq g_2 \leq \ldots$, all the functions g_n are measurable, Theorem 2.13. Taking the $\liminf$ in the above inequality we have

$$\liminf_{n\to\infty} \int_X g_n d\mu \leq \liminf_{n\to\infty} \int_X f_n d\mu.$$

By Theorem 3.5

$$\liminf_{n\to\infty} \int_X g_n d\mu = \int_X \lim_{n\to\infty} g_n d\mu.$$

We conclude since $\lim_{n\to\infty} g_n = \liminf_{k\to\infty} f_k$. $\square$

Exercise 3.10. *Construct a sequence of measurable functions $f_n : \mathbb{R} \to [0,\infty]$ such that the inequality in Fatou's Lemma, Corollary 3.9, is sharp.*

Hint: the conditions in Theorem 3.26 below must be violated.

We can also extend Lemma 3.6

Theorem 3.11. *Let f be a nonnegative μ-measurable function on X. Denote by Σ the μ-measurable sets. The map $\tilde{\nu} : \Sigma \to [0,\infty]$ given by*

$$\tilde{\nu}(E) := \int_E f\, d\mu \quad \forall E \ \mu\text{-measurable}$$

is a premeasure, cf. Definition 1.39.

As before, we denote by $f \llcorner \mu$ its Carathéodory-Hahn Extension (concatenation of f and μ).

Moreover for every measurable function $g : X \to [0,\infty]$ we have

$$\int_X g\, d\,(f \llcorner \mu) = \int_X g f\, d\mu. \tag{3.3}$$

Proof. As in the proof of Lemma 3.6 we only need to show that whenever $E_1, E_2, \ldots \in \Sigma$ are disjoint μ-measurable sets and $E = \bigcup_{n=1}^{\infty} E_n$ then

$$f \llcorner \mu(E) = \sum_n f \llcorner \mu(E_n).$$

Now observe that

$$\chi_E f(x) = \sum_{n=1}^{\infty} \chi_{E_n} f(x) \quad \forall x \in X.$$

Thus

$$(f \llcorner \mu)(E) = \int_X \chi_E f d\mu = \int_X \sum_{n=1}^{\infty} \chi_{E_n} f.$$

By Lebesgue Monotone Convergence theorem, Theorem 3.5, we can take out the sum, and have

$$(f \llcorner \mu)(E) = \sum_{n=1}^{\infty} \int_X \chi_{E_n} f = \sum_{n=1}^{\infty} (f \llcorner \mu)(E_n).$$

For (3.3) observe that it holds by definition if $g = \chi_E$ whenever E is μ-measurable. Since any nonnegative μ-measurable function can be approximated by simple functions, Theorem 2.14, we conclude the (3.3) again from Lebesgue Monotone Convergence theorem, Theorem 3.5. $\square$

3.1. **L^p-spaces and Lebesgue dominated convergence theorem.** We now want to discuss not only nonnegative measurable functions $f : X \to [0, \infty]$, but want to integrate general measurable functions $f : X \to \bar{\mathbb{R}}$.

Definition 3.12. Let (X, Σ, μ) be a measure space and $f : X \to \bar{\mathbb{R}}$.

- We say that f is *μ-integrable* if $|f| : X \to [0, \infty]$[11] if

  $$\|f\|_{L^1(X,\mu)} := \int_X |f| < \infty.$$

  If f is integrable we write

  $$\int_X f d\mu := \int_X f_+ d\mu - \int_X f_- d\mu$$

  where as usual $f_+ := \max\{f, 0\}$ and $f_- = -\min\{f, 0\}$. Observe that f integrable assures us that $\int_X f_+ d\mu < \infty$ and $\int_X f_- d\mu < \infty$ so there is no issue with $\infty - \infty$.
- More generally for $p \in (1, \infty)$ set

  $$\|f\|_{L^p(X,\mu)} := \left(\int_X |f(x)|^p \, dx \right)^{\frac{1}{p}}.$$

  If $\|f\|_{L^p(X,\mu)} < \infty$ we say that $f \in L^p(X, \mu)$.
- We also have the *L^2-scalar product* or the *L^2-pairing*

  $$\langle f, g \rangle := \int_X f(x) \, g(x) \, d\mu$$

  Let us remark (although we make no substantial use of this) that for *complex* functions $f, g : X \to \mathbb{C}$ the scalar product is

  $$\langle f, g \rangle := \int_X f(x) \, \overline{g(x)} \, d\mu$$

  where $\overline{g}$ is the complex conjugation.

  Clearly

  $$\langle f, f \rangle = \|f\|^2_{L^2(X,\mu)}.$$
- For $p = \infty$ we define the *essential supremum* (often still denoted by sup)

  $$\|f\|_{L^\infty(X,d\mu)} := \sup_X |f| := \inf \{\Lambda \in [0, \infty] : \quad \mu\{x \in X : |f(x)| > \Lambda\} = 0\}$$

  In words, $\sup_X |f|$ is the smallest number Λ such that the superlevel set of $\{f > \Lambda\}$ has zero measure.
- If $X \subset \mathbb{R}^n$ and μ is the Lebesgue measure we drop the μ and simpy write $L^p(X)$.

Exercise 3.13. *Let $p \in [1, \infty]$, $f, g \in L^p(X, \mu)$ and $\lambda \in \mathbb{R}$. Show that*

(1) (Homogeneity) $\|\lambda f\|_{L^p(X,\mu)} = |\lambda| \|f\|_{L^P(X,\mu)}$,
(2) (Minkowski-inequality) $\|f + g\|_{L^p(X,\mu)} \leq \|f\|_{L^p(X,\mu)} + \|g\|_{L^p(X,\mu)}$,

[11]which is μ-measurable

(3) *(Hölder-inequality)* $\|fg\|_{L^p(X,\mu)} \leq \|f\|_{L^q(X,\mu)}\|g\|_{L^r(X,\mu)}$, *where* $q, r \in [1,\infty]$ *are such that*

$$\frac{1}{p} = \frac{1}{q} + \frac{1}{r}.$$

(Here $\frac{1}{\infty} := 0$*).*

In particular,

$$\|f\|_{L^p(X,\mu)} \leq \mu(X)^{\frac{1}{r}} \|f\|_{L^q(X,\mu)}$$

(4) *(generalized Hölder's inequality)* $\|f_1 \cdot \ldots \cdot f_n\|_{L^p(X,\mu)} \leq \|f_1\|_{L^{q_1}(X,\mu)} \cdot \ldots \|f_n\|_{L^{q_n}(X,\mu)}$, *where* $q_i, p \in [1,\infty]$ *are such that*

$$\frac{1}{p} = \sum_{i=1}^{n} \frac{1}{q_i}$$

(5) *(Jensen) Let* $\mu(X) < \infty$, $f : X \to (-\infty, \infty)$ *and* $\varphi : \mathbb{R} \to \mathbb{R}$ *convex, then*

$$\varphi\left(\mu(X)^{-1} \int_X f(x) d\mu\right) \leq \mu(X)^{-1} \int_X \varphi(f(x)) d\mu$$

Hint: Adv.Calc or Wikipedia

Exercise 3.14. *Let* $p \in [1,\infty)$. *Assume* $(f_k)_{k\in\mathbb{N}} \subset L^p(X,\mu)$ *with*

$$\sup_{k\in\mathbb{N}} \|f_k\|_{L^p(X,\mu)} < \infty.$$

Assume moreover there exists $f : X \to \bar{\mathbb{R}}$ *such that* $f(x) = \lim_{k\to\infty} f_k(x)$ *for* μ*-a.e.* $x \in X$.

Using Fatou's Lemma, Corollary 3.9, show that $f \in L^p(X,\mu)$ *and we have*

$$\|f\|_{L^p(X,\mu)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(X,\mu)}.$$

Exercise 3.15. (1) *If* f *is continuous and* μ *is the Lebesgue measure on an open set* Ω *show that the essential supremum coincides with the usual supremum. (Show also this is not the case if* Ω *contains e.g. isolated points)*

(2) *Give an example of measurable* f *where essential supremum and supremum is not the same.*

Exercise 3.16. *Assume* $\mu(X) < \infty$ *and* $f : X \to \mathbb{R}$ *is* μ*-measurable. Show that*

$$\lim_{p\to\infty} \left(\mu(X)^{-1} \int_X |f|^p d\mu\right)^{\frac{1}{p}} = \|f\|_{L^\infty(X)}.$$

Exercise 3.17. *We have*

(1)
$$|f(x)| \leq \|f\|_{L^\infty(\mu)} \quad \mu\text{-a.e. } x.$$

(2) *If* $|f(x)| \leq \Lambda$ *for* μ*-a.e.* x, *then* $\|f\|_{L^\infty} \leq \Lambda$.

Exercise 3.18. *Let $1 \leq p \leq r \leq q \leq \infty$ and assume $f \in L^p(X)$ and $f \in L^q(X)$. Then $f \in L^r(X)$ and we have*

$$\|f\|^r_{L^r(X)} \leq \|f\|^q_{L^q(X)} + \|f\|^p_{L^p(X)}.$$

Hint: Let $A := \{x : f(x) \leq 1\}$. Then $f(x) = f(x)\chi_A(x) + f(x)\chi_{X\setminus A}(x)$

So the collection of μ-measurable functions $f : X \to \bar{\mathbb{R}}$ such that $f \in L^p(X, \mu)$ is a *linear space*: if $f, g : X \to \bar{\mathbb{R}}$ satisfy $\|f\|_{L^p(X,\mu)}$, $\|g\|_{L^p(X,\mu)} < \infty$ then for any $\lambda_1, \lambda_2 \in \mathbb{R}$ we have $\|\lambda_1 f + \lambda_2 g\|_{L^p(X,\mu)} < \infty$.

$\|\cdot\|_{L^p(X,\mu)}$ is a *pseudonorm* on this linear space. Recall that for a linear space L a map $\|\cdot\| : L \to [0, \infty)$ is a *norm* iff

(1) $\|f\| = 0$ if and only if $f = 0$
(2) $\|f + g\| \leq \|f\| + \|g\|$ for all $f, g \in L$
(3) $\|\lambda f\| = |\lambda|\|f\|$ for all $f \in L$, $\lambda \in \mathbb{R}$.

If the first property (1) fails, i.e. if there are $f \in L$ such that $\|f\| = 0$ but $f \neq 0$, then $\|\cdot\|$ is a *pseudonorm*.

We don't want to work with pseudonorms. (Don't worry, this looks more complicated than it is). We will instead work with classes of functions. From now on we say that two μ-measurable functions $f, g : X \to \overline{\mathbb{R}}$ are equal

$$f = g \quad \Leftrightarrow f(x) = g(x) \quad \mu\text{-a.e.}.$$

(check: this is an equivalence relation).

In particular we will often define a function only in $X \backslash N$ where $\mu(N) = 0$.

The integral does not see differences on zero-sets.

Exercise 3.19. *Let $f, g : X \to \bar{\mathbb{R}}$ be μ-measurable and $f = g$ in the above sense. Then*

$$\int_X f d\mu = \int_X g d\mu$$

in the following sense:

- *$f \in L^1(X)$ if and only if $g \in L^1(X)$.*
- *If $f \in L^1(X)$ (or equivalently $g \in L^1(X)$) then*

$$\int_X f d\mu = \int_X g d\mu.$$

Then we define

$$L^p(X, \mu) := \left\{ f : X \to \bar{\mathbb{R}} : \mu\text{-measurable and } \|f\|_{L^p(X,\mu)} < \infty \right\} / =$$

That is: an element in $L^p(X, \mu)$ is not a single function, but a class of functions (two functions belong to the same class if they only differ on a zero set). We often brush over this fact by saying f is a function (but actually meaning f's class). We don't even need to define a function everywhere, defining it outside a zero-set is enough. A specific function f in a class $[f]$ is called a *representative*.

While most of the time we don't bother whether we talk about a specific representative f or its class $[f]$ – sometimes we do care. For example if f is a continuous representative of its class $[f]$. The good news is that there cannot be two different continuous representatives of a class $[f]$, Exercise 4.51 for the Lebesgue measure.

Example 3.20.
- The function
$$f(x) = \frac{1}{|x|}$$
is $\mathcal{L}^n$-measurable in $\mathbb{R}^n$ and L^p-integrable in $B(0,1) \subset \mathbb{R}^n$ if $p < n$.
Clearly the $f(0)$ is not defined, but $\{0\}$ is a $\mathcal{L}^n$-zeroset, so it does not matter. A proper representative of $\frac{1}{|x|}$ could be
$$g(x) := \begin{cases} \frac{1}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$
or
$$g(x) := \begin{cases} \frac{1}{|x|} & x \neq 0 \\ \infty & x = 0 \end{cases}$$
or
$$g(x) := \begin{cases} \frac{1}{|x|} & x \neq 0 \\ \infty & x = 0 \\ 736 & x = (7,0,0,\ldots,0). \end{cases}$$
- Let $f \in C^0(\Omega)$ for some bounded open set Ω. Take any set A with $\mu(A) = 0$, and set
$$g := f + \chi_A.$$
Then $g = f$ μ-a.e. – that is not every representative of a class containing a continuous representative is continuous. Indeed, quite the opposite: there is at most one continuous representative in each class if μ is the Lebesgue measure, Exercise 4.51.

Lemma 3.21. *Assume $f \in L^p(X,\mu)$ for some $p \in [1,\infty]$. Then (for any representative of f), $|f(x)| < \infty$ μ-a.e.*

Proof. Set $A := \{x : |f(x)| = \infty\}$. Then we have (assume $p < \infty$, exercise for $p = \infty$!)
$$\mu(A) \cdot \infty = \int_A |f(x)|^p d\mu \leq \int_X |f(x)|^p d\mu = \|f\|^p_{L^p(X,\mu)} < \infty.$$
The only way this is possible is if $\mu(A) = 0$. □

Lemma 3.22. *(1) Suppose $f : X \to [0,\infty]$ is measurable and E is measurable. Then $\int_E f d\mu = 0$ if and only if $f = 0$ a.e. in E.*

(2) *Suppose* $f \in L^1(X,\mu)$ *and* $\int_E f d\mu = 0$ *for all* E *measurable. Then* $f = 0$ *a.e. in* X

Proof. (1) Set $A_n := \{x \in E : f(x) \geq \frac{1}{n}\}$. Then[12]

$$\frac{1}{n}\mu(A_n) = \int_{A_n} \frac{1}{n} d\mu \leq \int_{A_n} f d\mu \overset{f\geq 0}{\leq} \int_E f d\mu = 0.$$

Thus $\mu(A_n) = 0$ for all $n \in \mathbb{N}$. Therefore the set $\{x \in E : f(x) > 0\} = \bigcup_{n=1}^{\infty} A_n$ has measure zero.

(2) Define $E = \{x \in X : f(x) \geq 0\}$. Then

$$0 = \int_E f d\mu.$$

By part (1) we conclude that $f = 0$ a.e. in E. Arguing the same way for $X \backslash E = \{f(x) < 0\}$ we conclude.

$\square$

Exercise 3.23. *For the classes of functions* $f \in L^p(X,\mu)$, $\|\cdot\|_{L^p}$ *is indeed a norm. (Hint: Use Lemma 3.22)*

Definition 3.24. Let $p \in [1,\infty]$ and $f : X \to \mathbb{R}^N$ μ-measurable. We say $f \in L^p(X,\mu,\mathbb{R}^n)$ (meaning there is a class of which f is a representative) if and only if $|f| \in L^p(X,\mu)$. We can also define the integral componentwise

$$\int_X f \ d\mu = \left(\int_X f^1 \ d\mu, \dots, \int_X f^n \ d\mu\right).$$

Lemma 3.25. *If* $f \in L^1(X,\mu,\mathbb{R}^n)$ *then*

$$\left|\int_X f \, d\mu\right| \leq \int_X |f| d\mu.$$

Proof. Since $\int_X f \, d\mu \in \mathbb{R}^N$ there exists[13] a vector $v \in \mathbb{R}^N$ with $|v| = 1$ such that

$$|\int_X f| = \langle v, \int_X f d\mu\rangle = \sum_{i=1}^n v^i \int_X f^i d\mu.$$

So we have

$$|\int_X f| = \int_X \sum_{i=1}^n v^i f^i d\mu = \int_X \langle v, f\rangle d\mu$$

Now pointwise (Cauchy-Schwarz)

$$|\langle v, f\rangle| \leq |f|,$$

which implies the claim. $\square$

[12] The first inequality will later be called the Chebycheff inequality, Lemma 3.48

[13] if $w \in \mathbb{R}^n$ is a vector then $v := \frac{w}{|w|}$ (if $|w| \neq 0$) or v with $|v| = 1$ if $w = 0$) satisfies $\langle v, w\rangle = |w|$

Theorem 3.26 (Lebesgue dominated convergence theorem)**.** *Suppose $(f_n)_{n\in\mathbb{N}}$ is a sequence of μ-measurable functions $f_n : X \to \overline{\mathbb{R}}$ and $f : X \to \overline{\mathbb{R}}$ a function with $f(x) = \lim_{n\to\infty} f_n(x)$ for μ-almost every $x \in X$*[14].

If there exists $g \in L^1(X, \mu)$ such that

$$|f_n(x)| \leq g(x) \quad \mu\text{-a.e. } x \in X,\ \forall n \in \mathbb{N}$$

then $f \in L^1(\mu)$ and

$$\lim_{n\to\infty} \|f - f_n\|_{L^1(X,\mu)} = 0,$$

in particular

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Proof. By setting f_n, f, g to be zero on a zeroset we can assume that all the "almost everywhere" above can be replaced by everywhere.

f as a pointwise limit of measurable functions f_n is measurable, and clearly $|f| \leq g$ everywhere in X. Thus $\int_X |f| d\mu \leq \int_X g d\mu < \infty$, i.e. $f \in L^1(X, \mu)$.

Moreover we have $2g - |f - f_n| \geq 0$, so we can apply Fatou's lemma, Corollary 3.9,

$$\begin{aligned}
\int_X 2g d\mu &= \int_X \lim_{n\to\infty} (2g - |f - f_n|)\, d\mu \\
&\leq \liminf_{n\to\infty} \int_X (2g - |f - f_n|)\, d\mu \\
&= \int_X 2g d\mu + \liminf_{n\to\infty} \left(-\int_X |f - f_n| d\mu\right) \\
&= \int_X 2g d\mu - \limsup_{n\to\infty} \left(\int_X |f - f_n| d\mu\right)
\end{aligned}$$

The integral $\int_X 2g d\mu < \infty$. Subtracting it form both sides of the above inequality we find

$$\limsup_{n\to\infty} \left(\int_X |f - f_n| d\mu\right) \leq 0.$$

Since the integral is nonnegative we conclude

$$\lim_{n\to\infty} \|f - f_n\|_{L^1(X,\mu)} = \lim_{n\to\infty} \left(\int_X |f - f_n| d\mu\right) = 0.$$

The last claim follows from Lemma 3.25,

$$\left|\int_X f_n d\mu - \int_X f d\mu\right| \leq \int_X |f_n - f|\, d\mu \xrightarrow{n\to\infty} 0.$$

□

The next theorem, while it is now not extremely difficult anymore, is one of the main advantages of the Lebesgue integral: it leads to *complete* L^p-spaces.

[14] observe without additional assumptions there is no hope of having $\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu$, Exercise 3.10

Theorem 3.27 (L^p is complete)**.** *Let $(f_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $L^p(X,\mu)$, that is*

$$\forall \varepsilon > 0 \quad \exists K \in \mathbb{N}: \quad \|f_k - f_\ell\|_{L^p} < \varepsilon \quad \forall k, \ell \geq K.$$

Then there exists $f \in L^p(X,\mu)$ such that

$$\|f_k - f\|_{L^p(X,\mu)} \xrightarrow{k\to\infty} 0.$$

Proof. By assumption, for each k there exists a number $\varphi(k)$ such that

$$\|f_\ell - f_j\|_{L^p} < 2^{-k} \quad \forall \ell, j \geq \varphi(k).$$

W.l.o.g. $\varphi(k) \leq \varphi(k+1)$.

We claim that

$$f(x) := \lim_{k\to\infty} f_{\varphi(k)}(x)$$

exists μ-a.e., and that this limit is actually an L^p-limit. To show that we write

$$f_{\varphi(k)} = \sum_{j=1}^{k-1} \left(f_{\varphi(j+1)} - f_{\varphi(j)} \right) + f_{\varphi(1)}.$$

Now define

$$G(x) := \sum_{j=1}^{\infty} \left| f_{\varphi(j+1)}(x) - f_{\varphi(j)}(x) \right|$$

As a limit of μ-measurable functions, G is μ-measurable.

Either by monotonce convergence theorem (if $p < \infty$ or by hand if $p = \infty$) we have

$$\|G\|_{L^p(X,\mu)} \leq \sum_{j=1}^{\infty} \|f_{\varphi(j+1)} - f_{\varphi(j)}\|_{L^p(X,\mu)} \leq \sum_{j=1}^{\infty} 2^{-j} < \infty.$$

That is $G \in L^p(X,\mu)$. Thus G is μ-a.e. finite, Lemma 3.21. Consequently, we have absolute convergence of the series

$$F(x) := \sum_{j=1}^{\infty} \left(f_{\varphi(j+1)}(x) - f_{\varphi(j)}(x) \right) \in \mathbb{R} \quad \mu\text{-a.e.},$$

so F is μ-measurable as limit of μ-measurable functions. And we also have

$$|F(x)| \leq G(x) \quad \mu\text{-a.e.}.$$

In particular, $\|F\|_{L^p(X,\mu)} < \infty$. Since the series is μ-a.e. absolute convergence we can set for μ-a.e. x

$$f(x) := \lim_{k\to\infty} f_{\varphi(k)}(x) = \lim_{k\to\infty} \sum_{j=1}^{k-1} \left(f_{\varphi(j+1)(x)} - f_{\varphi(j)} \right)(x) + f_{\varphi(1)}(x) = F(x)) + f_{\varphi(1)}(x),$$

which exists and is a finite number μ-a.e.. As a sum of L^p-functions $f \in L^p(X,\mu)$.

Using again the telescoping sum we have

$$|f - f_{\varphi(\ell)}| = \sum_{k=\ell}^{\infty} |f_{\varphi(k+1)} - f_{\varphi(k)}| \leq F \quad \forall \ell \in \mathbb{N}$$

Then we can apply dominated convergence, Theorem 3.26, and have

$$\lim_{\ell\to\infty} \|f - f_{\varphi(\ell)}\|_{L^p(X)} = \|\lim_{\ell\to\infty} \left(f - f_{\varphi(\ell)}\right)\|_{L^p(X)} = 0.$$

That is $(f_{\varphi(\ell)})_\ell$ converges in L^p to f.

This is only a subsequence, but with the Cauchy-condition we can conclude. Fix $\varepsilon > 0$, then there must be some K such that

$$\|f_k - f_\ell\|_{L^p} \leq \varepsilon \quad \forall k, \ell \geq K.$$

Now let $k \geq K$ and pick any ℓ such that $\varphi(\ell) \geq K$. Then

$$\|f_k - f\|_{L^p} \leq \|f_k - f_{\varphi(\ell)}\|_{L^p} + \|f_{\varphi(\ell)} - f\|_{L^p} \leq \varepsilon + \|f_{\varphi(\ell)} - f\|_{L^p}.$$

Taking $\ell \to \infty$ we find

$$\|f_k - f\|_{L^p} \leq \varepsilon \quad \forall k \geq K.$$

Thus f_k converges to f in L^p. □

Theorem 3.28. *Let X be a locally compact metric space and μ a Radon measure on X. Then the class of compactly supported continuous functions $C_c(X)$ is dense in $L^p(\mu)$ for all $1 \leq p < \infty$.*

Exercise 3.29. *Show that Theorem 3.28 is false for $p = \infty$ (hint: L^∞ is uniform convergence. What do we know about continuity under uniform convergence?)*

For the proof of Theorem 3.28 we need the following Lemma:

Lemma 3.30. *Let S be the class of finite, measurable, simple functions s on X such that*

$$\mu(\{x : s(x) \neq 0\}) < \infty.$$

If $1 \leq p < \infty$, then S is dense in $L^p(\mu)$.

Proof. Clearly $S \subset L^p(\mu)$. If $f \in L^p(\mu)$ and $f \geq 0$, let s_n be a sequence of simple functions such that $0 \leq s_n \leq f$, $s_n \to f$ pointwise. Since $s_n \in L^p$ it easily follows that $s_n \in S$. Now inequality $0 \leq |f - s_n|^p \leq f^p$ and the dominated convergence theorem implies that $s_n \to f$ in L^p. In the general case we write $f = (u^+ - u^-)$ and apply the above argument to each of the functions u^+ , u^- separately. □

Proof of Theorem 3.28. According to Lemma 3.30 it suffices to prove that the characteristic function of a set of finite measure can be approximated in L^p by compactly supported continuous functions. Let E be a measurable set of finite measure. Given $\varepsilon > 0$ let $K \subset E$ be a compact set such that $\mu(E \setminus K) < (\varepsilon/2)^p$ – this is possible by Theorem 1.68. Let U be an open set such that $K \subset U$, $\overline{U}$ is compact and $\mu(U \setminus K) < (\varepsilon/2)^p$. Finally let

$\varphi \in C_c(X)$ be such that $\operatorname{supp} \varphi \subset U$, $0 \leq \varphi \leq 1$, $\varphi(x) = 1$ for $x \in K$ – for example if we set $d := \frac{1}{2}\text{dist}\,(K, X \setminus U)$ we can choose $\varphi(x) := \frac{(d - \text{dist}\,(x,K))_+}{d}$. We then have

$$\begin{aligned}\|\chi_E - \varphi\|_{L^p} = &\left(\int_{X\setminus K} |\chi_E - \varphi|^p d\mu\right)^{\frac{1}{p}} \leq \left(\int_{X\setminus K} |\chi_E|^p d\mu\right)^{\frac{1}{p}} + \left(\int_{X\setminus K} |\varphi|^p d\mu\right)^{\frac{1}{p}}\\ \leq &\mu(E \setminus K)^{1/p} + \mu(U \setminus K)^{1/p} < \varepsilon.\end{aligned}$$

□

If $X = \mathbb{R}^n$ (or any smooth manifold) it is easy to change $C^0_c(X)$ to $C^\infty_c(X)$ in Theorem 3.28, by choosing a suitable smoother version of φ (see also Exercise 4.34).

Definition 3.31. A set $A \subset X$ for a metric space X is called *dense* if for any $x \in X$ and any $\varepsilon > 0$ there exists $a \in A$ with $d(a, x) < \varepsilon$.

A metric space X is *separable* if there is a *countable* dense set.

Theorem 3.32. *Let $\Omega \subset \mathbb{R}^n$ be open, μ a nonzero-Radon measure, and $1 \leq p < \infty$, then*

(1) $L^p(\Omega, \mu)$ is separable
(2) $C^0_c(\Omega)$ is dense in $L^p(\Omega, \mu)$, where

$$C^0_c(\Omega) := \{f \in C^0(\Omega) : \quad \operatorname{supp} f \subset \Omega\},$$

and supp f *denotes the* support of f*,*

$$\operatorname{supp} f = \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

(3) $C^\infty_c(\Omega)$ is dense in $L^p(\Omega, \mu)$, where

$$C^\infty_c(\Omega) := \{f \in C^\infty(\Omega) : \quad \operatorname{supp} f \subset \Omega\}.$$

Exercise 3.33. *Show Theorem 3.32.*

Hint: for (1) use Theorem 3.28 and Stone-Weierstrass. (2) and (3) can be proven similarly to Theorem 3.28.

By Stone-Weierstrass $C^0(\overline{\Omega})$ is a separable space if Ω is a bounded open set, moreover $C^0(\overline{\Omega})$ is a closed subset of $L^\infty(\Omega)$ with Lebesgue measure. It is a strict subset, because $L^\infty(\Omega)$ is not separable (so C^0_c is not dense, and the Theorem 3.32 does not hold for $p = \infty$).

Exercise 3.34. *Let $\Omega = [0, 1]$, $\mu = \mathcal{L}^1$. Set $f_t := \chi_{[0,t]}$, $0 < t \leq 1$. Show that*

(1) $\|f_t - f_s\|_{L^\infty} = 1$ whenever $s \neq t$.
(2) $(f_t)_{t\in[0,1]}$ is uncountable
(3) Thus L^∞ is not separable.

3.2. **Lebesgue integral vs Riemann integral.** How do we actually compute the Lebesgue integral? Well, if we really want to compute the Lebesgue integral chances are we can just use the Riemann integral. They coincide on Riemann-integrable functions.

Theorem 3.35. *Let $f : -\infty < a < b < \infty \to \mathbb{R}$ be Riemann-integrable (in the classical Darboux sense. No improper Riemann integrability allowed!). Then f is Lebesgue integrable, i.e. $f \in L^1([a,b], \mathcal{L}^1)$ and we have*

$$\mathrm{Ri} - \int_{[a,b]} f(x)dx = \int_{[a,b]} f(x) d\mathcal{L}^1(x).$$

Proof. W.l.o.g. $a = 0$ and $b = 1$. Clearly Riemann integral and Lebesgue integral coincide on constant functions.

Since f is Riemann-integrable it is bounded, and by adding a constant to f (changing the integrals equally), we may assume w.l.o.g. that f is nonnegative.

The Riemann-Lebesgue theorem states that a function f is Riemann integrable if and only if f is bounded, and f is continuous outside of a set N of zero Lebesgue measure.

Then $f : [0,1]\backslash N \to [0,\infty)$ is continuous and thus $\mathcal{L}^1$-measurable (because f^{-1} maps open sets to open sets). But then $f : [0,1] \to [0,\infty)$ is also measurable since

$$f^{-1}(A) = \left(f\Big|_{[0,1]\backslash N} \right)^{-1} (A) \cup \underbrace{(f^{-1}(A) \cap N)}_{\text{zeroset}}.$$

Thus f is measurable and nonnegative, and since it is also bounded it is indeed $\mathcal{L}^1$-integrable with

$$\int_{[0,1]} f(x) d\mathcal{L}^1(x) \leq \sup_{[0,1]} f \; \mathcal{L}^1([0,1]) < \infty.$$

We still need to show that the two integrals coincide.

Fix $\varepsilon > 0$. By the Darboux sum definition there exists a partition $0 = x_0 < x_1 < \ldots < x_N = 1$ such that for $m_i := \inf_{z \in [x_{i-1}, x_i]} f(z)$ and $M_i := \sup_{z \in [x_{i-1}, x_i]} f(z)$ we have

$$\sum_{i=1}^N (x_i - x_{i-1}) M_i - \varepsilon \leq \mathrm{Ri} \int_{[0,1]} f(x)dx \leq \sum_{i=1}^N (x_i - x_{i-1}) m_i + \varepsilon$$

Set

$$g(x) := m_i \quad \text{where } i \text{ is such that } x \in (x_{i-1}, x_i)$$

and

$$G(x) := M_i \quad \text{where } i \text{ is such that } x \in (x_{i-1}, x_i)$$

(observe g_i and G_i are a.e. defined only, but that is enough).

Since g_i and G_i are step functions, we have

$$\int_{[0,1]} g d\mathcal{L}^1(x) = \sum_{i=1}^N \mathcal{L}^1\left((x_{i-1}, x_i)\right) m_i = \sum_{i=1}^N (x_i - x_{i-1}) m_i$$

and

$$\int_{[0,1]} G d\mathcal{L}^1(x) = \sum_{i=1}^{N} \mathcal{L}^1\left((x_{i-1}, x_i)\right) M_i = \sum_{i=1}^{N} (x_i - x_{i-1}) M_i.$$

That is we have

$$\int_{[0,1]} G d\mathcal{L}^1(x) - \varepsilon \leq \mathrm{Ri} \int_{[0,1]} f(x)dx \leq \int_{[0,1]} g d\mathcal{L}^1(x) + \varepsilon$$

Now observe that by definition $g \leq f \leq G$ a.e. . So we end up with

$$\int_{[0,1]} f d\mathcal{L}^1(x) - \varepsilon \leq \mathrm{Ri} \int_{[0,1]} f(x)dx \leq \int_{[0,1]} f d\mathcal{L}^1(x) + \varepsilon$$

This implies that

$$\left| \int_{[0,1]} f d\mathcal{L}^1(x) - \mathrm{Ri} \int_{[0,1]} f(x)dx \right| < \varepsilon.$$

This holds for any $\varepsilon > 0$ and by letting $\varepsilon \to 0$ we obtain

$$\int_{[0,1]} f d\mathcal{L}^1(x) = \mathrm{Ri} \int_{[0,1]} f(x)dx.$$

□

So we conclude that for many situations the Lebesgue-integral is just the Riemann integral. But the Lebesgue integral allows often for more functions (they have usually no practical meaning, but are theoretically important).

Example 3.36. The Dirichlet function

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \not\in \mathbb{Q} \end{cases}$$

is not Riemann-integrable, however since $f(x) = 0$ $\mathcal{L}^1$-a.e. (exercise!) we have that f is Lebesgue integrable and $\int f(x)d\mathcal{L}^1 = 0$.

Since the n-dimensional Riemann integral is just (by Fubini's theorem) several 1-dimension integrals one can extend this theorem to any dimensions and domains as long as they are Riemann-measurable domains (once we have the Fubini theorem for measures, Section 4).

One can also conclude that the improper integrals for *nonnegative* functions coincide (by dominated convergence theorem for the Lebesgue integral and definition for the Riemann integral).

However this is not true for general improper integrals:

Exercise 3.37. *Show that* $\mathrm{Ri} - \int_{\mathbb{R}} \frac{\sin(x)}{x}\, dx$ *is finite, but show that* $\frac{\sin(x)}{x}$ *is not Lebesgue-integrable.*

So we conclude that Lebesgue integral and Riemann integral are pretty much the same from a practical point of view (i.e. “I want to compute the integral” – indeed the Riemann integral is easier to compute e.g. numerically).

The Lebesgue integral however are better theoretical spaces, the main advantage being that many limits of integrable functions are integrable (under mild assumptions – because even the *a.e.*-limit stays measurable), whereas for Riemann-integral you need something like uniform convergence. So one of the main reasons of doing all of this is the completeness above Theorem 3.27. Another reason is that we have much more variety for measures (and thus integrals), we can integrate w.r.t Hausdorff measure etc.

3.3. **Theorems of Lusin and Egorov.** We have seen in Theorem 2.14 that measurable functions are “infinite step functions”. But more is true. Measurable functions are continuous on large parts of their domain (Lusin). Moroever pointwise convergence of measurable functions is uniform convergence on a large set (Egorov).

Theorem 3.38 (Egorov)**.** *Let X be a metric space and μ any Radon measure. Suppose that $\mu(X) < \infty$.*

Let $f_k, f : X \to \bar{\mathbb{R}}$ μ-measurable and μ-a.e. finite[15]*, i.e. $|f(x)| < \infty$ for μ-a.e. $x \in X$.*

Moreover for μ-a.e. $x \in X$ assume

$$\lim_{k\to\infty} f_k(x) = f(x).$$

Then for any $\delta > 0$ there exists a compact set $F \subset X$ such that

$$\mu(X\backslash F) < \delta$$

and we have uniform convergence $f_k\Big|_F \xrightarrow{k\to\infty} f\Big|_F$, namely

$$\sup_{x\in F} |f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

Remark 3.39.
- Taking $f_k := \chi_{B(0,k)}$ which converges pointwise to $f \equiv 1$ in $\mathbb{R}^n$, however there is no compact set $\mathcal{L}^n(\mathbb{R}^n\backslash F) < \infty$. So the assumption $\mu(X) < \infty$ is in general needed necessary.
- The result of Theorem 3.38 is not possible for $\delta = 0$ (i.e. uniform convergence outside of a zero-set may not be true), take the usual $f_k(x) = x^k$ in $[0, 1]$ which does converge pointwise but not uniformly in $[0, 1]$, nor any dense subset of $[0, 1]$, Exercise 1.80.

[15]Alternatively we could assume f_k or f in $L^1(X, \mu)$. Indeed, if $f \in L^1(X, \mu)$ then f is μ-a.e. finite, Lemma 3.21.

Proof of Theorem 3.38. Fix $\delta > 0$. For each $i, j \in \mathbb{N}$ set

$$C_{i,j} := \bigcup_{k=j}^{\infty} \{x \in X : |f_k(x) - f(x)| > 2^{-i}\}.$$

Observe that pointwise convergence f_k to f a.e. $\mu(\bigcap_j C_{i,j}) = 0$ for each i. Indeed, if $x \in \bigcup_{k=j}^{\infty}\{z \in X : |f_k(z) - f(z)| > 2^{-i}\}$ for every j, then there are infinitely many k such that $x \in \{z \in X : |f_k(z) - f(z)| > 2^{-i}\}$, so there is a subsequence $(f_{k_\ell})_{\ell \in \mathbb{N}}$ such that

$$|f_{k_\ell}(x) - f(x)| \geq 2^{-i}.$$

But by a.e. convergence of f_k this means that x belongs to a zero set.

The pointwise limit of measurable functions is measurable, so f and also $C_{i,j}$ is measurable, Theorem 2.13.

Moroever we have $C_{i,j+1} \subset C_{i,j}$.

Since $\mu(X) < \infty$ we can apply Theorem 1.33 (3) and have

$$\lim_{j \to \infty} \mu(C_{i,j}) = \mu(\bigcap_{j=1}^{\infty} C_{i,j}) = 0.$$

In particular for any $i \in \mathbb{N}$ there exists $N(i)$ such that

$$\mu(C_{i,N(i)}) < \delta 2^{-i-1}.$$

Set $A := X \backslash \bigcup_{i=1}^{\infty} C_{i,N(i)}$ then we have

$$\mu(X \backslash A) \leq \sum_{i=1}^{\infty} \mu(C_{i,N(i)}) < \frac{\delta}{2}.$$

Also, by definition of $C(i, N(i))$ we have

$$\sup_{x \in A} |f_k(x) - f(x)| \leq \sup_{x \notin C(i,N(i))} |f_k(x) - f(x)| \leq 2^{-i}, \quad \forall k \geq N(i).$$

That is: f_k converges uniformly to f in A!

Since A is measurable and $\mu(A) \leq \mu(X) < \infty$, we apply Theorem 1.68 and find some compact $F \subset X$ such that

$$\mu(A) \leq \mu(F) + \frac{\delta}{2},$$

i.e. (as a Radon measure, μ is Borel-measure, so F is μ-measurable)

$$\mu(A \backslash F) \leq \frac{\delta}{4}.$$

In particular we have

$$\mu(X \backslash F) \leq \mu(X \backslash A) + \mu(A \backslash F) < \delta.$$

Since f_k uniformly converges to f in A we can conclude. $\square$

Theorem 3.40 (Lusin)**.** *Let X be a metric space and μ any Radon measure. Suppose that $\mu(X) < \infty$.*

Let $f : X \to \bar{\mathbb{R}}$ be a μ-measurable and μ-a.e. finite, i.e. $|f(x)| < \infty$ μ-a.e.

Then there exists $\delta > 0$ and $F \subset X$ compact such that

- $\mu(X \backslash F) < \delta$
- $f\Big|_F : F \to \mathbb{R}$ *is continuous.*

Remark 3.41. Observe that

$$f\Big|_F : F \to \mathbb{R} \quad \text{is continuous}$$

and

$$f : X \to \mathbb{R} \quad \text{is continuous in F}$$

are *not* the same.

For example if we take $X = [0,1]$, $f := \chi_{[0,1]\backslash\mathbb{Q}}$ then f is nowhere continuous in $[0,1]$. However $f : \Big|_{[0,1]\backslash\mathbb{Q}}$ is constant and thus continuous as map $[0,1]\backslash\mathbb{Q} \to \mathbb{R}$.

In particular this example also shows (for the Lebesgue measure) that Theorem 3.40 may not hold for $\delta = 0$ (i.e. outside of a zero-set) because in general one cannot find a *compact* F for $\delta = 0$.

Proof of Theorem 3.40. Step 1 We show the statement for simple functions

$$g := \sum_{i=1}^{I} b_i \chi_{B_i},$$

where B_i are measurable, pairwise disjoint sets with $X = \bigcup_{i=1}^{I} B_i$, and $b_i \in \mathbb{R}$.

Fix $\delta > 0$. By Theorem 1.68 for each B_i (since $\mu(B_i) < \infty$) there exists a compact $F_i \subset B_i$ such that

$$\mu(B_i) \leq \mu(F_i) + \delta 2^{-i}.$$

In particular F_i is μ-measurable since F_i is a Borel set and μ is Borel measure. Thus,

$$\mu(B_i \backslash F_i) < \delta 2^{-i} \quad 1 \leq i \leq I.$$

Since $(B_i)_i$ are pairwise disjoint, so are the sets $(F_i)_i$. Since each F_i is compact this implies $\operatorname{dist}(F_i, F_j) > 0$ for each $i \neq j$. So if we set

$$F := \bigcup_{i=1}^{I} F_i \subset X$$

then we have that $g\Big|_F$ is locally constant (namely: for each $x_0 \in F$ there exists a radius $\rho = \rho(x_0) > 0$ such that $g\Big|_{F\cap B(x_0,\rho)}$ is constant). This implies $g\Big|_F$ is continuous as a map $F \to \mathbb{R}$ (and thus in particular bounded).

Moreover,

$$\mu(X\backslash F) = \mu(\bigcup_{i=1}^{I}(B_i\backslash F_i)) \leq \sum_{i=1}^{I}\mu(B_i\backslash F_i) < \delta.$$

Step 2 Since $|f(x)| < \infty$ for μ-a.e. x, up to changing f on a μ-zero set (which does not influence the result, because we can always pass to a smaller (compact) set if needed to get rid of the zero-set). Now if we write $f = f^+ - f^-$ and apply Theorem 2.14 to f_+ and f_- then we have

$$f(x) = \sum_{\ell=1}^{\infty}\frac{1}{\ell}\chi_{A_\ell^+}(x) - \sum_{\ell=1}^{\infty}\frac{1}{\ell}\chi_{A_\ell^-}(x) \quad \text{everywhere in } X$$

In particular by writing

$$f_k(x) = \sum_{\ell=k}^{\infty}\frac{1}{\ell}\chi_{A_\ell^+}(x) - \sum_{\ell=1}^{k}\frac{1}{\ell}\chi_{A_\ell^-}(x).$$

Since these are finitely many sets, we can make them disjoint, that is we can write

$$f_k(x) = \sum_{i=1}^{I_k} b_i\chi_{B_i}$$

where $(B_i)_{i=1}^{\infty}$ are pairwise disjoint, and by adding $X\backslash\bigcup_{i=1}^{I_k}B_i$ (which still has finite measure) we have that f_k satisfies the conditions of Step 1.

So fix again $\delta > 0$, apply Step 1 to f_k and we find compact sets $F_k \subset X$ such that for each $k \in \mathbb{N}$.

$$\mu(X\backslash F_k) < \delta 2^{-k-1}$$

and

$$f_k\Big|_{F_k} : F_k \to \mathbb{R} \quad \text{is continuous}$$

We furthermore find a compact set $F_0 \subset X$ from Egorov's theorem, Theorem 3.38, such that

$$\mu(X\backslash F_0) < \frac{\delta}{2}$$

and

$$\sup_{x\in F_0}|f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

Now set $F := \bigcap_{k=0}^{\infty} F_k \subset X$. Then F is compact and

$$\mu(X\backslash F) \leq \mu(\bigcup_{k=0}^{\infty}(X\backslash F_k)) \leq \sum_{k=0}^{\infty}\mu(X\backslash F_k) < \delta.$$

Since $F \subset F_0$ we have uniform convergence of $f_k\Big|_F$ to $f\Big|_F$. Since each $f_k\Big|_F$ is continuous, its uniform limit $f\Big|_F : F \to \mathbb{R}$ is continuous. $\square$

Corollary 3.42 (Lusin in Euclidean space)**.** *Let $E \subset \mathbb{R}^n$ be Lebesgue-measurable. Then a function $f : E \to \mathbb{R}$ (i.e. f is pointwise finite!) is Lebesgue measurable if and only for any $\varepsilon > 0$ there is a closed*[16] *set $F \subset E$ such that $f\Big|_F : F \to \mathbb{R}$ is continuous and $\mathcal{L}^n(E\backslash F) < \varepsilon$.*

Proof. $\underline{\Rightarrow}$ Fix $\delta \in (0,1)$.

Let $A_0 := \overline{B(0,1)}$ and $A_k := \overline{B(0,k\delta)}\backslash B(0,k(1-2^{-k-2})\delta)$, $k \geq 1$.

Observe that this way

$$\mathcal{L}^n(A_k \cap A_\ell) \begin{cases} = 0 & |\ell - k| \geq 2 \\ \leq C(n)2^{-k}k\delta^n & |\ell - k| = 1. \end{cases}$$

Apply Theorem 3.40 to each A_k. We then find a compact $\tilde{F}_k \subset A_k$ such that $\mathcal{L}^n(A_k\backslash F_k) < 2^{-k}\delta$ and $f\Big|_{\tilde{F}_k}$ is continuous. Setting $F_k := \tilde{F}_k \cap \overline{B(0,k\delta(1-2^{-k-1})}$ we have that the (F_k) are pairwise disjoint compact sets, so $f\Big|_{\bigcup_k F_k}$ is continuous and still we have

$$\mathcal{L}^n(A_k\backslash F_k) \leq \mathcal{L}^n(A_k\backslash \tilde{F}_k) + \mathcal{L}^n(\overline{B(0,k\delta)}\backslash\overline{B(0,k(1-2^{-k-1}\delta)})) \leq C(n)\delta 2^{-k}(1+k).$$

In particular $\mathbb{R}^n = \bigcup_{k=1}^\infty A_k$ so if we set $\tilde{F} := \bigcup_{k=1}^\infty F_k$ then $f\Big|_{\tilde{F}}$ is continuous and

$$\mathcal{L}^n(\mathbb{R}^n\backslash\tilde{F}) \leq \sum_{k=1}^\infty \mathcal{L}^n(A_k\backslash F_k) \leq C(n)\delta.$$

However $\tilde{F}$ may not be closed. But in view of Theorem 1.68 there exists an open set $G \supset \mathbb{R}^n\backslash\tilde{F}$ such that

$$\mathcal{L}^n(G) \leq 2\mathcal{L}^n(\mathbb{R}^n\backslash\tilde{F}).$$

Set $F := \mathbb{R}^n\backslash G$ then F is closed and $F \subset \tilde{F}$. So $f\Big|_F$ is continuous and

$$\mathcal{L}^n(\mathbb{R}^n\backslash F) = \mathcal{L}^n(G) \leq 2C(n)\delta.$$

If we chose δ so that $2C(n)\delta < \varepsilon$ we can conclude.

$\underline{\Leftarrow}$ f is measurable iff $E_a := \{x : f(x) \geq \alpha\}$ is measurable for each α, (cf. Lemma 2.4, and observe that f is pointwise finite)

[16] it clearly can't be compact. Namely $\mathcal{L}^n(\mathbb{R}^n\backslash F) = \infty$ for any bounded set F

Let F be a closed set such that $f\Big|_F$ is continuous and $\mathcal{L}^n(E\backslash F) < \varepsilon$. Then the set

$$F'_a := \{x \in F : \quad f\Big|_F(x) \geq a\} = E_a \cap F$$

is closed (this follows from the continuity of f). Thus

$$\mathcal{L}^n(E_a \backslash F'_a) = \mathcal{L}^n(E_a \backslash F) \leq \mathcal{L}^n(E \backslash F) < \varepsilon.$$

Since this argument works for any $\varepsilon > 0$, from Theorem 1.70(3) we obtain that E_a is measurable.

□

3.4. **Convergence in measure.** For sequences of μ-measurable functions $f_k : X \to \mathbb{R}$, we have learned about pointwise convergence

$$f_k(x) \xrightarrow{k\to\infty} f(x) \quad \forall x$$

slighly weaker, μ-almost everywhere convergence

$$f_k(x) \xrightarrow{k\to\infty} f(x) \quad \mu\text{-a.e. } x,$$

then (way stronger) than a.e. uniform convergence

$$\|f_k - f\|_{L^\infty(X,\mu)} = \sup_{x\in X} |f_k(x) - f(x)| \xrightarrow{k\to\infty} 0.$$

Recall that from now on we consider sup to be the essential sup $\operatorname{ess\,sup}$.

From the definition of L^p-spaces we also have L^p-convergence,

$$\|f_k - f\|_{L^p(X,\mu)} = \left(\int_X |f_k(x) - f(x)|^p d\mu\right)^{\frac{1}{p}} \xrightarrow{k\to\infty} 0.$$

Uniform convergence is then very similar to L^∞-convergence if the functions in question are continuous and the domains are open sets.

Now we introduce a weaker notion of convergence than pointwise, namely *convergence in measure*.

Definition 3.43. Let (X, μ) be a metric measure space. We say that a sequence of μ-measurable functions $(f_k)_{k\in\mathbb{N}}$, $f_k : X \to \mathbb{R}$ *converges in measure* to $f : X \to \mathbb{R}$ is for every $\varepsilon > 0$

$$\lim_{n\to\infty} \mu\left(\{x \in X : \ |f_n(x) - f(x)| > \varepsilon\}\right) = 0.$$

Some people (us including) will write sometimes $f_n \xrightarrow{\mu} f$. As usual we assume that the functions f_n and f are defined a.e.

Theorem 3.44 (Lebesgue)**.** *Assume $\mu(X) < \infty$ and a sequence of measurable functions $f_n : X \to \mathbb{R}$ converges to $f : X \to \mathbb{R}$ μ-almost everywhere. Then $f_n \xrightarrow{\mu} f$.*

Proof. Fix any $\varepsilon > 0$. Set

$$E_i := \{x \in X : |f_i(x) - f(x)| \geq \varepsilon\}.$$

The sequence of sets $A_n := \bigcup_{i=n}^{\infty} E_i$ is decreasing sequence of measurable sets, all with finite measure, and hence, Theorem 1.33,

$$\mu\left(\bigcup_{i=n}^{\infty} E_i\right) \xrightarrow{n\to\infty} \mu\left(\bigcap_{n=1}^{\infty}\bigcup_{i=n}^{\infty} E_i\right) = 0.$$

The last set has measure zero, because if $x \in \bigcup_{i=n}^{\infty} E_i$ for every n, then x belongs to infinitely many E_i's, so there is a subsequence f_{n_i} such that

$$|f_{n_i}(x) - f(x)| \geq \varepsilon$$

which is only true on a set of measure zero, by the a.e. convergence.

Hence

$$\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) = \mu(E_n) \leq \mu(\bigcup_{i=n}^{\infty} E_i) \xrightarrow{n\to\infty} 0,$$

which proves convergence in measure. $\square$

The converse of Theorem 3.44 does not hold (see Exercise 3.47 below), but it holds up to subsequence.

Theorem 3.45 (Riesz)**.** *If $f_n \xrightarrow{\mu} f$ then there is a subsequence $(f_{n_i})_{i\in\mathbb{N}}$ such that $f_{n_i} \to f$ a.e.*

Observe that here $\mu(X) = \infty$ is permissible.

Proof. For every $i \in \mathbb{N}$ there exist n_i such that

$$\mu\left(\{x : \quad |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}\right) \leq 2^{-i}.$$

We can assume that $n_1 < n_2 < n_3$. Let

$$F_k := X \backslash \bigcup_{i=k}^{\infty}\{x : |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}$$

and

$$F := \bigcup_{k=1}^{\infty} F_k.$$

Then

$$\mu(X\backslash F_k) \leq \sum_{i=k}^{\infty} 2^{-i} = 2^{1-k}.$$

and thus

$$\mu(X \setminus F) \leq \limsup_{k\to\infty} \mu(X \setminus F_k) \leq \limsup_{k\to\infty} 2^{1-k} = 0.$$

We claim that

$$f_{n_i}(x) \xrightarrow{i\to\infty} f(x) \quad \text{for any } x \in F. \tag{3.4}$$

Since $X \setminus F$ is a μ-zeroset we can conclude once we have proven this.

It suffices to prove (3.4) for each F_k, and actually we will show that for each F_k we have uniform convergence.

Indeed, if $x \in F_k$ then x then $x \not\in \bigcup_{i=k}^{\infty} \{x : |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}$. That is

$$|f_{n_j}(x) - f(x)| \leq \frac{1}{j} \quad \forall j \geq k, \quad \forall x \in F_k$$

This is uniform convergence in F_k. □

Remark 3.46. We actually proved not only convergence a.e. but also uniform convergence on subsets of X whose complement has arbitrary small measure. Note that Lebesgue's theorem Theorem 3.44 combined with this stronger conclusion of Riesz' theorem Theorem 3.45 implies Egorov's theorem for sequences of real valued functions, Theorem 3.38.

One does indeed need to pass to a subsequence in Theorem 3.45, as the following example shows.

Exercise 3.47. *Let $f(x) := \chi_{[0,1]}$. For each $n \in \mathbb{N}$ there exists exactly one $m \in \mathbb{N}$ and $k \in \{0, \ldots, 2^m - 1\}$ such that $n = 2^m + k$. Set*

$$f_n(x) := f(2^m x - k).$$

Show

- *f_n converges to 0 in $L^2([0,1])$*
- *$f_n(x)$ does not converge to 0 for any $x \in (0,1)$.*

3.5. **L^p-convergence and weak L^p.** Next we want to consider convergence in L^p-norm. For this we first show

Lemma 3.48 (Chebyshev's inequality)**.** *Let $f : X \to \mathbb{R}$ be μ-integrable. Then for any $\lambda > 0$, any $p \in [1,\infty)$*

$$\mu(\{x : f(x) > \lambda\}) \leq \frac{1}{\lambda^p} \int_{\{x: f(x)>\lambda\}} |f(x)|^p \, dx,$$

and in particular

$$\mu(\{x : f(x) > \lambda\}) \leq \frac{1}{\lambda^p} \|f\|^p_{L^p(X)}.$$

Proof. Since f is measurable the set $\{x : f(x) > \lambda\}$ is measurable and thus

$$\begin{aligned}\mu(\{x : f(x) > \lambda\}) =& \int_{\{x:f(x)>\lambda\}} d\mu\\ =&\lambda^{-p} \int_{\{x:f(x)>\lambda\}} \lambda^p d\mu\\ \leq&\lambda^{-p} \int_{\{x:f(x)>\lambda\}} |f|^p d\mu.\end{aligned}$$

□

As a slight digression, Chebyshev's inequality leads to a space very similar to L^p, but weaker – hence called *weak L^p-space*, denoted by $L^{(p,\infty)}$.

Definition 3.49 (weak L^p)**.** We say that a μ-measurable $f : X \to \mathbb{R}$ belongs to weak L^p, $f \in L^{(p,\infty)}(X, \mu)$ if there exists $\Lambda \geq 0$ such that for all $\lambda > 0$

$$\mu(\{x : f(x) > \lambda\}) \leq \frac{1}{\lambda^p}\Lambda^p.$$

The minimal value of Λ such that the above inequality holds is denoted by $\|f\|_{L^{(p,\infty)}(X,\mu)}$.

$$\|f\|_{L^{(p,\infty)}(X,\mu)} := \sup_{\lambda>0} \lambda \left(\mu(\{x : f(x) > \lambda\})\right)^{\frac{1}{p}}$$

One can check that $\|f\|_{L^{(p,\infty)}(X,\mu)}$ is not a true norm (triangle inequality is only true up to a multiplicative constant).

Chebyshev's inequality implies that $\|f\|_{L^{(p,\infty)}} \leq \|f\|_{L^p}$. The other direction is not true.

Exercise 3.50. *Show that for $\sigma \in (0, n]$ the function $f(x) = |x|^{-\sigma}$ belongs to $L^{(\frac{n}{\sigma},\infty)}(\mathbb{R}^n)$. Show that $f \notin L^q(\mathbb{R}^n)$ for any $q \in [1, \infty]$.*

Weak L^p-spaces become very important in Harmonic Analysis, they can also be generalized to so-called Lorentz spaces $L^{(p,q)}$.

As a corollary of Chebyshev inequality we get

Theorem 3.51. *Let $p \in [1, \infty]$ and assume that $f_k \xrightarrow{k\to\infty} f \in L^p(X, \mu)$, i.e.*

$$\|f_k - f\|_{L^p(X,\mu)} \xrightarrow{k\to\infty} 0.$$

Then

- $f_k \xrightarrow{\mu} f$
- *there exists a subsequence $k_i \to \infty$such that $f_{k_i} \xrightarrow{i\to\infty} f$ μ-a.e.*

Proof of Theorem 3.51. The second claim follows from the first one from Riesz theorem, Theorem 3.45. For the first claim observe that by Lemma 3.48

$$\mu\{|f_k - f| > \varepsilon\} \leq \varepsilon^{-p}\|f_k - f\|^p_{L^p} \xrightarrow{k\to\infty} 0.$$

□

Here is another criterion for convergence

Theorem 3.52. *Let $p \in [1,\infty)$ and assume $f, f_k \in L^p(X,\mu)$ for some metric space X. Then the following are equivalent*

(1) $\|f_k - f\|_{L^p(X)} \xrightarrow{k\to\infty} 0$
(2) $f_k \xrightarrow{\mu} f$ *and* $\|f_k\|_{L^p(X)} \xrightarrow{k\to\infty} \|f\|_{L^p(X)}$.

Exercise 3.53. *Show Theorem 3.52 is false for $p = \infty$ (hint: continuous functions)*

Proof of Theorem 3.52. (1) ⇒ (2): The convergence in measure follows from Theorem 3.51. Moreover we have by reverse triangle inequality

$$\left|\|f_k\|_{L^p(X)} - \|f\|_{L^p(X)}\right| \leq \|f_k - f\|_{L^p(X)} \xrightarrow{k\to\infty} 0.$$

(2) ⇒ (1): First assume that $f_k(x) \xrightarrow{k\to\infty} f(x)$ for μ-a.e. x.

Observe that by convexity (here we need $1 \leq p < \infty$)

$$2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p \geq 0.$$

Moreover, by pointwise convergence

$$\liminf_{k\to\infty} 2^{p-1}|f_k(x)|^p + 2^{p-1}|f(x)|^p - |f_k(x) - f|^p = 2 * 2^{p-1}|f(x)| \quad \mu\text{-a.e. } x$$

By assumption (2) we have

$$\liminf_{k\to\infty} \int_\Omega |f_k|^p = \int_\Omega |f|^p,$$

so by Fatou's Lemma, Corollary 3.9,

$$\begin{aligned}
&2^p \int_\Omega |f|^p d\mu - \limsup_{k\to\infty} \int_\Omega |f_k - f|^p d\mu \\
=&\liminf_{k\to\infty} \int_\Omega \left(2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p\right) d\mu \\
\geq&\int_\Omega \liminf_{k\to\infty} \left(2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p\right) d\mu \\
=&2^p \int_\Omega |f|^p d\mu
\end{aligned}$$

Subtracting the (finite!) $2^p \int_\Omega |f|^p d\mu$ from both sides of this inequality we find

$$\limsup_{k\to\infty} \|f_k - f\|^p_{L^p(\Omega,\mu)} = \limsup_{k\to\infty} \int_\Omega |f_k - f|^p d\mu = 0.$$

This proves (1) under the assumption that $f_k(x) \xrightarrow{k\to\infty} f(x)$ for μ-a.e. x and $\|f_k\|_{L^p(X)} \xrightarrow{k\to\infty} \|f\|_{L^p(X)}$.

Assume now $f_k \xrightarrow{\mu} f$ and $\|f_k\|_{L^p(X)} \xrightarrow{k\to\infty} \|f\|_{L^p(X)}$. Assume to the contrary that there exists some subsequence f_{k_i} such that

$$\|f_{k_i} - f\|_{L^p(X)} > \varepsilon \quad \forall i \in \mathbb{N}. \tag{3.5}$$

By Theorem 3.45 we can find a subsubsequence $(f_{k_{i_j}})_j$ such that $f_{k_{i_j}}(x) \xrightarrow{j\to\infty} f(x)$ for μ-a.e. x. But then we have by the above arguments that

$$\|f_{k_{i_j}} - f\|_{L^p(X)} \xrightarrow{j\to\infty} 0.$$

This contradicts (3.5). □

We will later discuss another way to decide when the convergence in measure implies L^p-convergence in Theorem 3.59 below, but first we need to talk about absolute continuity.

3.6. **Absolute continuity.** Let $f : X \to [0,\infty)$ μ-integrable. For μ-measurable $A \subset \Omega$ set

$$\nu(A) := \mu\llcorner f = \int_A f d\mu.$$

Cf. Theorem 3.11. Observe that then $\mu(A) = 0$ implies $\nu(A) = 0$ – an effect which we call *absolute continuity*.

Definition 3.54. Let μ, ν be measures on X such that

- any μ-measurable set is also ν-measurable
- if $\mu(A) = 0$ then $\nu(A) = 0$.

Then we say that ν is *absolutely continuous* with respect to μ, and write $\nu \ll \mu$.

Lemma 3.55. *Let μ and ν be Radon measures on $\mathbb{R}^n$*[17]

Then

(1) Assume that for all $A \subset \mathbb{R}^n$ it holds that if $\mu(A) = 0$ then also $\nu(A) = 0$. Then any μ-measurable set is also ν-measurable.
(2) In particular, $\mu \ll \nu$ if and only if for all $A \subset \mathbb{R}^n$ if $\mu(A) = 0$ then also $\nu(A) = 0$.

Proof. We use Theorem 1.70: any μ-measurable set A can be written as $A = B \cup N$ where B is Borell and $\mu(N) = 0$. So $\nu(N) = 0$ and consequently $A = B \cup N$ is ν-measurable.

□

The notion of continuity is justified by the following (important on its own) *absolute continuity of the integral*, and Exercise 3.57.

[17] easily extendable to more general sets that satisfy the assumptions of Theorem 1.70

Theorem 3.56. *Let $f \in L^1(\mu)$. Then for every $\varepsilon > 0$ there is $\delta > 0$ such that $\int_E |f| d\mu < \varepsilon$ whenever E is measurable and $\mu(E) < \delta$.*

Proof. We argue by contradiction. If the claim was false, there would exist $\varepsilon_0 > 0$ and μ-measurable sets E_n such that $\mu(E_n) < 2^{-n}$ but $\int_{E_n} |f| d\mu \geq \varepsilon_0$.

The sequence of sets $A_k := \bigcup_{n=k}^{\infty} E_n$ is decreasing and

$$\mu(\bigcap_{k=1}^{\infty} A_k) = 0,$$

since

$$\mu(\bigcap_{k=1}^{\infty} A_k) \leq \mu(A_n) \leq \sum_{i=n}^{\infty} \mu(E_i) \leq 2^{1-n} \xrightarrow{n\to\infty} 0.$$

Now

$$|f| \llcorner \mu(E) = \int_E |f| d\mu$$

is a measure, Theorem 3.11, so we have

$$\int_{A_k} |f| d\mu = |f| \llcorner \mu(A_k) \xrightarrow{k\to\infty} |f| \llcorner \mu(\bigcap_{k=1}^{\infty} A_k) = \int_{\bigcap_{k=1}^{\infty} A_k} |f| d\mu = 0,$$

since $\mu(\bigcap_{k=1}^{\infty} A_k) = 0$.

On the other hand from the choice of E_n we have

$$\int_{A_k} |f| d\mu \geq \int_{E_k} |f| d\mu \geq \varepsilon_0,$$

a contradiction. □

More generally we can reformulate Theorem 3.56 in the following way

Exercise 3.57. *Let $\mu, \nu : 2^X \to [0, \infty]$ be two measures with $\nu \ll \mu$ and assume $\nu(X) < \infty$. Then*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 : \quad \forall A \subset X \ \mu\text{-measurable:} \mu(A) < \delta \Rightarrow \nu(A) < \varepsilon.$$

3.7. **Vitali's convergence theorem.** We know that if f_k converges to f in L^p then the convergence is also in measure, Theorem 3.51. Vitali's convergence theorem is a characterization when the other direction is true, i.e. when convergence in measure (or in view of Theorem 3.44 a.e. convergence if on a sets of finite measure) implies L^p-convergence.

Definition 3.58. A family $\mathcal{F}$ of μ-integrable functions $f : X \to \bar{\mathbb{R}}$ is said to have *uniformly absolutely continuous integrals* if

$$\forall \varepsilon > 0 \, \exists \delta > 0 : \quad \forall f \in \mathcal{F} : \quad \forall A \subset X : \ \mu\text{-measurable} : \quad \mu(A) < \delta \Rightarrow \int_A |f| d\mu < \varepsilon.$$

(Compare this notion with Arzela-Ascoli!)

Theorem 3.59 (Vitali convergence theorem)**.** *Let X be a metric space and μ a Radon measure. Assume $\mu(X) < \infty$ and let $(f_k)_{k\in\mathbb{N}}$ be a sequence in $L^1(X, d\mu)$ and $f \in L^1(X, d\mu)$. Then the following are equivalent*

(1) $f_k \xrightarrow{\mu} f$ *and* $(f_k)_{k\in\mathbb{N}}$ *has uniformly absolutely continuous integrals*
(2) $\lim_{k\to\infty} \int_X |f_k - f| d\mu = 0.$

Example 3.60. The assumption $\mu(X) < \infty$ is needed in general. E.g. take $\mu = \mathcal{L}^n$, $X = \mathbb{R}^n$, $f_k = k^{-n}\chi_{B(0,k)}$. Then $f_k \xrightarrow{\mu} 0$ (because the convergence is a.e.). $(f_k)_k$ has uniformly absolutely continuous integrals, but $\int_{\mathbb{R}^n} |f_k| d\mu = c > 0$ with a constant $c = \mathcal{L}^n(B(0,1))$ independent of k, but $\int 0 = 0$.

Proof of Theorem 3.59. (2) ⇒ (1): If $\|f_k - f\|_{L^1} \xrightarrow{k\to\infty} 0$ then $f_k \xrightarrow{\mu} f$ by Theorem 3.51. To get the uniform absolute continuous integral property fix $\varepsilon > 0$.

In view of L^1-convergence there exists a $K \in \mathbb{N}$ such that

$$\|f_k - f\|_{L^1(X)} \le \frac{\varepsilon}{2} \quad \forall k \ge K.$$

Moreover, by absolute continuty of the integral, Theorem 3.56, there exists a $\delta > 0$ such that any μ-measurable set $A \subset X$ with $\mu(A) < \delta$ satisfies

$$\max_{k\in\{1,\dots,K\}} \int_A |f_k| + \int_A |f| < \frac{\varepsilon}{2}.$$

Since

$$\int_A |f_k| \le \int_A |f_k - f| + \int_A |f| d\mu \le \|f_k - f\|_{L^1(X)} + \int_A |f| d\mu,$$

we find that for all $k \ge K$ we have

$$\int_A |f_k| < \varepsilon.$$

The same holds also for $k \in \{1, \dots, K\}$, so we have shown the uniform absolute continuity of $(f_k)_k$.

(1) ⇒ (2): Let $f_k \xrightarrow{\mu} f$ with uniform absolute continuous integrals.

Assume by contradiction that

$$\limsup_{k\to\infty} \int_\Omega |f_k - f| d\mu > 0.$$

By passing to a subsequence we can assume w.l.o.g.

$$\liminf_{k\to\infty} \int_\Omega |f_k - f| d\mu > 0. \tag{3.6}$$

Passing yet to another subsequence, applying Theorem 3.45, we may assume that

$$f_k \xrightarrow{k\to\infty} f \quad \text{a.e. in } X.$$

Fix $\varepsilon > 0$. Since f_k are uniformly absolutely continuous and f is integrable we can find $\delta > 0$ such that for any μ-measurable $A \subset X$ with $\mu(A) < \delta$ we have

$$\int_A |f| d\mu < \varepsilon, \quad \int_A |f_k| d\mu < \varepsilon \quad \forall k \in \mathbb{N}.$$

For this δ we can apply Egorov's theorem, Theorem 3.38, and find a compact set $F \subset X$ with $\mu(X \backslash F) < \delta$ and

$$\sup_{x \in F} |f_k(x) - f(x)| \xrightarrow{k \to \infty} 0.$$

In particular there must be $K \in \mathbb{N}$ such that

$$\sup_{k \geq K} \sup_{x \in F} |f_k(x) - f(x)| < \frac{\varepsilon}{2\mu(X)}$$

Consequently, for any $k \geq k_0$ we have

$$\begin{aligned}
\int_X |f_k - f| d\mu &= \int_{X \backslash F} |f_k - f| d\mu + \int_F |f_k - f| d\mu \\
&\leq \int_{X \backslash F} |f_k| + |f| d\mu + \int_F \sup_{k \geq K} \sup_{x \in F} |f_k - f| d\mu \\
&\leq \varepsilon + \varepsilon + \varepsilon,
\end{aligned}$$

the first two estimates are because of absolute continuity, the last one because of uniform convergence.

In particular we conclude that

$$\liminf_{k \to \infty} \int_X |f_k - f| d\mu \leq 3\varepsilon,$$

since $\varepsilon > 0$ was arbitary this is a contradiction to (3.6). □

Exercise 3.61. *Formulate and prove a Vitali-type condition for L^p-convergence.*

In the calculus of variations Vitali's theorem can be used for minimization.

Definition 3.62. Let $\Omega \subset \mathbb{R}^n$. A map $F = F(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ is a *Carathéodory-function* if

- for all $y \in \mathbb{R}$ the function $x \mapsto F(x, y)$ is μ-measurable
- for μ-a.e. $x \in \Omega$ the function $y \mapsto F(x, y)$ is continuous.

Exercise 3.63. *If F is a Carathéodory function and $u : \Omega \to \mathbb{R}$ is μ-measurable then $f(x) := F(x, u(x))$ is μ-measurable.*

Example 3.64. Let $\Omega \subset \mathbb{R}^n$ with $\mu(\Omega) < \infty$, $F = F(x, y) : \Omega \times \mathbb{R} \to \mathbb{R}$ a *Carathéodory-function* as above. Assume moreover that F has a *p-growth* for $p \in [1, \infty)$, namely

$$|F(x, y)| \leq C(1 + |y|^p) \quad \mu\text{-a.e. } x \in \Omega,\ y \in \mathbb{R}.$$

Let furthermore $u, u_k : \Omega \to \mathbb{R}$ be μ-measurable functions with
$$\int_\Omega |u|^p d\mu < \infty$$
$$\int_\Omega |u_k|^p d\mu < \infty \quad \forall k$$
$$\int_\Omega |u_k - u|^p d\mu \xrightarrow{k\to\infty} 0.$$
Then
$$\int_\Omega F(x, u_k(x))d\mu \xrightarrow{k\to\infty} \int_\Omega F(x, u(x))d\mu$$

(This is much easier if F is also Lipschitz or Hölder continuous in y but we assume only continuity (and $u_k(\Omega)$ is not a k-uniformly bounded set!) We also cannot use dominated convergence, since we do not have a dominating function.

Solution. Set
$$f(x) = F(x, u(x)), \quad f_k(x) := F(x, u_k(x)).$$
In view of Exercise 3.63 f and f_k are μ-measurable.

Also
$$\int_\Omega |f_k(x)|d\mu \le C \int_\Omega (1 + |u_k(x)|^p)d\mu < \infty,$$
so f_k (and similarly f) are μ-integrable.

We want to argue with Theorem 3.59 and thus need to check its assumptions. $\underline{f_k \xrightarrow{\mu} f}$.

Assume this is not the case, then there exists some $c > 0$ such that
$$\limsup_{k\to\infty} \mu(\{x : |f_k - f| > \varepsilon\}) \ge c > 0$$
By the definition of lim sup there must be a subsequence f_{k_i} such that
$$\liminf_{i\to\infty} \mu(\{x : |f_{k_i} - f| > \varepsilon\}) \ge c > 0. \tag{3.7}$$
We will find that this leads to a contradiction.

Since u_{k_i} converges in L^p to u by Chebyshev inequality, more precisely by Theorem 3.51, we have $u_{k_i} \xrightarrow{\mu} u$.

Since $u_{k_i} \xrightarrow{\mu} u$ there exists a further subsequence $u_{k_{i_j}}$ such that $u_{k_{i_j}} \xrightarrow{j\to\infty} u$ μ-a.e. in Ω, see Theorem 3.45. Since F is continuous in the second entry we see that $f_{k_{i_j}} \xrightarrow{j\to\infty} f$ μ-a.e.

Since μ-a.e. convergence implies convergence in measure, Theorem 3.44 we found that $f_{k_{i_j}} \xrightarrow{\mu} f$ – but this implies
$$\lim_{j\to\infty} \mu(\{x : |f_{k_{i_j}} - f| > \varepsilon\}) = 0,$$
a contradiction to (3.7).

So indeed $f_k \xrightarrow{\mu} f$.

$\int |f_k| d\mu$ is uniformly absolutely continuous

Fix $\varepsilon > 0$. Since u_k L^p-converges to u by Vitali's theorem, Theorem 3.59, we have that $\int |u_k|^p$ is uniformly absolutely continuous, i.e. there exists $\delta > 0$ (and w.l.o.g. $\delta < \frac{\varepsilon}{2}$) such that for all μ-measurable sets A

$$\mu(A) < \infty \Rightarrow \int_A |u|^p d\mu + \sup_k \int_A |u_k|^p d\mu < \frac{\varepsilon}{2}.$$

Then we have

$$\int_A f_k d\mu \leq \int_A 1 + |u_k|^p d\mu = \mu(A) + \int_A |u_k|^p d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves uniform absolute continuity for $\int_A f_k d\mu$.

□

4. Product Measures, Multiple Integrals – Fubini's theorem

Fubini's theorem is essentially saying that if we want to integrate on a cube $[0,1]^2$ then we can write this as an integral on $[0,1] \times [0,1]$,

$$\int_{[0,1]^2} f(x,y) dx dy = \int_{[0,1]} \int_{[0,1]} f(x,y) dx dy.$$

Since we now work with more abstract measures, we first need to discuss what is dx dy...

Recall that the cartesian product of two spaces X and Y is given by

$$X \times Y = \{(x,y) : \quad x \in X : \quad y \in Y\}.$$

When measuring a set $A \times B$ it is easy to think that we somehow should multiply $\mu(A)\nu(B)$. But not all sets $S \subset X \times Y$ are of the form $A \times B$. What do we do? We take the best cover!

Definition 4.1 (Product measures)**.** Let $\mu : 2^X \to [0,\infty]$ and $\nu : 2^Y \to [0,\infty]$ be two measures. The *product measure* $\mu \times \nu : 2^{X \times Y} \to [0,\infty]$ for $S \subset X \times Y$ is defined as

$$(\mu \times \nu)(S) := \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) : \quad S \subset \bigcup_{i=1}^{\infty} A_i \times B_i; \quad A_i \subset X \ \mu\text{-measurable}, B_i \subset Y \ \nu\text{-measurable} \right\}$$

In the (most relevant) case of Lebesgue measure one can see now easily that $\mathcal{L}^k \times \mathcal{L}^\ell$ on $\mathbb{R}^k \times \mathbb{R}^\ell$ is indeeed $\mathcal{L}^{k+\ell}$ on $\mathbb{R}^{k+\ell}$.

Theorem 4.2 (Fubini's theorem)**.** *Let μ, ν be Radon measures on metric spaces X and Y.*

(1) If $A \subset X$ is μ-measurable, $B \subset Y$ is ν-measurable then $A \times B$ is $\mu \times \nu$-measurable and we have

$$(\mu \times \nu)(A \times B) = \mu(A)\nu(B).$$

(2) *Let* $S \subset X \times Y$ *be* $\mu \times \nu$*-measurable and* $\mu \times \nu(S) < \infty$*. Define for* $y \in Y$ *the set* $S_y \in X$ *as*
$$S_y := \{x : \quad (x,y) \in S\}$$
then S_y *is* μ*-measurable for* ν*-a.e.* y*. Moreover the map*
$$y \mapsto \mu(S_y) = \int_X \chi_S(x,y) d\mu(x)$$
is ν*-integrable and*
$$(\mu \times \nu)(S) = \int_Y \mu(S_y) d\nu(y) = \int_Y \left(\int_X \chi_S(x,y) d\mu(x)\right) d\nu(y).$$
Similarly for S_x*.*

(3) $\mu \times \nu$ *is a Radon measure*

(4) *Is* $f : X \times Y \to \bar{\mathbb{R}}$ $\mu \times \nu$*-integrable then*
$$y \mapsto \int_X f(x,y) d\mu(x)$$
is integrable w.r.t ν*, and*
$$x \mapsto \int_Y f(x,y) d\nu(y)$$
is integrable w.r.t. μ *and we have*
$$\int_{X \times Y} f \; d(\mu \times \nu) = \int_X \left(\int_Y f(x,y) \; d\nu(y)\right) d\mu(x) = \int_Y \left(\int_X f(x,y) \; d\mu(x)\right) d\nu(y).$$

This is a very technical statement, and we shall not go through the proof here (see [Evans and Gariepy, 2015, Theorem 1.22] for this). More instructive is to look at warnings:

Example 4.3.
- A set $S \subset X \times Y$ such that
$$x \mapsto \chi_S(x,y) \quad \text{is } \mu\text{-measurable for } \nu\text{-a.e. } y$$
and
$$y \mapsto \mu(S_y) \equiv \int_X \chi_S(x,y) d\mu(x) \quad \text{is } \nu\text{-measurable}$$
may *not* be measurable.

Indeed let $X = Y = \mathbb{R}$ and $\mu = \nu = \mathcal{L}^1$ and take $A \subset [0,1]$ the non-measurable Vitali set. Set
$$S := \{(x,y) : \quad |x - \chi_A(y)| < \frac{1}{2}, \quad y \in [0,1]\}.$$
We have S_y either $|x| < \frac{1}{2}$ or $|x-1| < \frac{1}{2}$ (both measurable), so S_y is μ-measurable for *all* y with $\mu(S_y) \equiv 1$.

However when $x > \frac{1}{2}$ then $S_x = A$. So S_x is not measurable for $\mathcal{L}^1$-a.e. x.

If S was $\mathcal{L}^2$-measurable then both S_x and S_y would need to be measurable (by Fubini's theorem).

- Similarly (take for example $f := \chi_S$ for the S from above) there is no reason that if $\int f d\mu$ exists and is integrable that then $f : X \times Y \to \mathbb{R}$ is integrable (or even measurable).

A slight variant of Fubini's theorem is the following *Tonelli's theorem*

Theorem 4.4 (Tonelli)**.** *Let $f : X \times Y \to [0, \infty]$ be $\mu \times \nu$-measurable, and assume that the iterated integral*

$$\int_X \left(\int_Y f(x,y) d\nu(y) \right) d\mu(x)$$

exists. Then f is $\nu \times \mu$ integrable and we have

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

For the Lebesgue measure and $f \in L^1(\mathbb{R}^n, \mathcal{L}^n)$ we often use as an application from Fubini's theorem

$$\int_{\mathbb{R}^n} f(x) d\mathcal{L}^n(x) = \int_{\mathbb{R}} \ldots \int_{\mathbb{R}} f(x_1, \ldots, x_n) d\mathcal{L}^1 x_1 \, d\mathcal{L}^1 x_2 \ldots d\mathcal{L}^1 x_n$$

This uses

Theorem 4.5. *The Lebesgue measure $\mathcal{L}^n$ on $\mathbb{R}^n$ satisfies*

$$\mathcal{L}^n = \mathcal{L}^k \times \mathcal{L}^{n-k} \quad \forall \kappa \in \{0, \ldots, n\}$$

In particular

$$\mathcal{L}^n = \mathcal{L}^1 \times \mathcal{L}^1 \times \ldots \times \mathcal{L}^1.$$

Proof. Both measures coincide on rectangles R (by construction), so for the application of Theorem 1.71 we only need that both measures are indeed Radon measures, which follows from Fubini's theorem Theorem 4.2. $\square$

A consequence of Fubini's theorem

Proposition 4.6 (Slicing)**.** *Assume $f \in L^1((0,R)^2)$ (take a representative and fix it). Then there exists $r \in (0,R)$ such that $x \mapsto f(x,r)$ belongs to $L^1((0,R))$ and moreover*

$$\int_{(0,R)} |f(x,r)| dx \leq \frac{1}{R} \int_{(0,R)^2} |f(x,y)| d(x,y).$$

More precisely, consider for $\lambda > 1$ the set $Y_\lambda \subset \mathbb{R}$

$$Y_\lambda := \left\{ r \in (0,R) : \int_{(0,R)} |f(x,r)| dx \leq \frac{\lambda}{R} \int_{(0,R)^2} |f(x,y)| d(x,y) \right\}.$$

then $\mathcal{L}^1(Y_\lambda) > (1 - \frac{1}{\lambda}) R$.

Proof. If
$$\int_{(0,R)^2} |f(x,y)|d(x,y) = 0$$
from Fubini's theorem we have
$$\int_{(0,R)} |f(x,r)|dx = 0 \quad \mathcal{L}^1\text{-a.e. } r \in (0,R),$$
and we can conclude.

So assume from now on
$$\int_{(0,R)^2} |f(x,y)|d(x,y) > 0.$$
Let
$$X_\lambda = \left\{ r \in (0,R) : \int_{(0,R)} |f(x,r)|dx > \frac{\lambda}{R} \int_{(0,R)^2} |f(x,y)|d(x,y) \right\}.$$
By Fubini's theorem X_λ is $\mathcal{L}^1$-measurable and $\mathcal{L}^1(X_\lambda) = \mathcal{L}^1((0,R)) - \mathcal{L}^1(Y_\lambda)$. For any $r \in X_\lambda$ we have
$$\frac{\lambda}{R} \int_{(0,R)^2} |f(x,y)|d(x,y) \leq \int_{(0,R)} |f(x,r)|dx$$
So integrating this in r we find
$$\mathcal{L}^1(X_\lambda)\frac{\lambda}{R} \int_{(0,R)^2} |f(x,y)|d(x,y) \leq \int_{X_\lambda} \int_{(0,R)} |f(x,r)|dxdr \leq \int_{(0,R)^2} |f(x,y)|dxdy$$
Dividing the integral on both sides we find
$$\mathcal{L}^1(X_\lambda) \leq \frac{R}{\lambda}.$$
Consequently, for any $\lambda > 1$
$$\mathcal{L}^1(Y_\lambda) \geq \mathcal{L}^1((0,R)) - \mathcal{L}^1(X_\lambda) = R - \frac{R}{\lambda} > 0$$
□

The measure estimate $\mathcal{L}^1(Y_\lambda) > (1-\frac{1}{\lambda})R$ in Proposition 4.6 is useful to show that several slicing properties hold on the same set simultaneously.

Exercise 4.7. *Show that there exists a uniform $\Lambda > 0$ such that the following holds:*

Take any (representative of) $f, g \in L^1((0,1)^2)$. There exists an $\mathcal{L}^1$-measurable set $Y \subset (0,1)$ with $\mathcal{L}^1(Y) > 0$ (depending on f and g) such that for each $r \in Y$ we have

(1) $x \mapsto f(x,r)$ belongs to $L^1((0,1))$
(2) $x \mapsto g(x,r)$ belongs to $L^1((0,1))$
(3) $\int_{(0,1)} |f(x,r)|dx \leq \Lambda \int_{(0,1)^2} |f(x,y)|d(x,y)$.
(4) $\int_{(0,1)} |g(x,r)|dx \leq \Lambda \int_{(0,1)^2} |f(x,y)|d(x,y)$.

Hint: *Use Proposition 4.6 for both f and g and obtain Y_f and Y_g where the above statements hold. To show that $\mathcal{L}^1(Y_f \cap Y_g) > 0$ use Exercise 4.8*

Exercise 4.8. *Let (X, μ) be a measure space with $\mu(X) < \infty$. Assume $A, B \subset X$ are μ-measurable with $\mu(A) > \frac{2}{3}\mu(X)$ and $\mu(B) > \frac{2}{3}\mu(X)$. Show that $\mu(A \cap B) \geq \frac{1}{6}\mu(X)$.*

Exercise 4.9. *Let $f \in L^1((0,1)^2)$ then for $\mathcal{L}^1$-a.e. $r \in (0,1)$ we have $f(r, \cdot) \in L^1(0,1)$.*

Hint: *Apply Proposition 4.6 to*

$$Y_k := \left\{ r \in (0,1) : \int_{(0,1)} |f(x,r)| dx \leq 2^k \int_{(0,1)^2} |f(x,y)| d(x,y) \right\}$$

and show $(0,1) \setminus \bigcup_k Y_k$ is a zero-set.

4.1. **Application: Interpolation between L^p-spaces – Marcienkiewicz interpolation theorem.** If $f \in L^{p_1}(\mathbb{R}^n) \cap L^{p_2}(\mathbb{R}^n)$ then $f \in L^p(\mathbb{R}^n)$ for $p \in (p_1, p_2)$, cf. Exercise 3.18.

So what if we have some information about a map T acting on $L^{p_1}(\mathbb{R}^n)$ and acting on $L^{p_2}(\mathbb{R}^n)$ – do we know something about how the maps acts on $L^p(\mathbb{R}^n)$. Under certain conditions yes.

As a first step observe that any L^p-map can be decomposed into an L^{p_1}-map and L^{p_2}-map.

Lemma 4.10. *Assume $1 \leq p_1 \leq p \leq p_2 \leq \infty$ and assume $f \in L^p(\mathbb{R}^n)$.*

For each fixed $\lambda > 0$ there exists f_1, f_2 with the following conditions

- $f_i \in L^{p_i}(\mathbb{R}^n)$, $i = 1, 2$.
- $f = f_1 + f_2$ *a.e. in* $\mathbb{R}^n$.

Moreover we have

$$\|f_i\|_{L^{p_1}(\mathbb{R}^n)} \leq \lambda^{1-\frac{p}{p_i}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{p_i}} \quad i = 1, 2.$$

f_i can be chosen explicitely:

$$f_1(x) := f(x)\chi_{|f|>\lambda} = \begin{cases} f(x) & \text{if } |f(x)| > \lambda \\ 0 & \text{if } |f(x)| \leq \lambda \end{cases}$$

and

$$f_2(x) := f(x)\chi_{|f|\leq\lambda} = \begin{cases} f(x) & \text{if } |f(x)| \leq \lambda \\ 0 & \text{if } |f(x)| > \lambda \end{cases}$$

Proof. We may assume $p_1 < p < p_2$ otherwise there is nothing to show.

Observe that since $\{f > \lambda\}$ etc. are measurable sets, so f_1 and f_2 are measurable functions, and we have $f = f_1 + f_2$ a.e..

It remains to compute (and here is where we use $p_1 < p < p_2$)

$$\begin{aligned}\|f_1\|_{L^{p_1}(\mathbb{R}^n)} =&\left(\int_{\{|f|>\lambda\}} |f(x)|^{p_1} dx\right)^{\frac{1}{p_1}}\\ =&\lambda^{1-\frac{p}{p_1}} \left(\int_{\{|f|>\lambda\}} \lambda^{p-p_1}|f(x)|^{p_1} dx\right)^{\frac{1}{p_1}}\\ \leq&\lambda^{1-\frac{p}{p_1}} \left(\int_{\{|f|>\lambda\}} |f(x)|^{p} dx\right)^{\frac{1}{p_1}}\\ \leq&\lambda^{1-\frac{p}{p_1}} \left(\int_{\mathbb{R}^n} |f(x)|^{p} dx\right)^{\frac{1}{p_1}}\\ =&\lambda^{1-\frac{p}{p_1}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{p_1}}.\end{aligned}$$

Similarly,

$$\begin{aligned}\|f_1\|_{L^{p_2}(\mathbb{R}^n)} =&\left(\int_{\{|f|\leq\lambda\}} |f(x)|^{p_2} dx\right)^{\frac{1}{p_2}}\\ \leq&\left(\int_{\{|f|\leq\lambda\}} \lambda^{p_2-p}|f(x)|^{p} dx\right)^{\frac{1}{p_2}}\\ \leq&\lambda^{1-\frac{p}{p_2}} \left(\int_{\mathbb{R}^n} |f(x)|^{p} dx\right)^{\frac{1}{p_2}}\\ =&\lambda^{1-\frac{p}{p_2}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{p_2}}.\end{aligned}$$

□

Now assume T is a linear operator from $L^{p_i}(\mathbb{R}^n)$ to $L^{q_i}(\mathbb{R}^n)$ for $i = 1, 2$. That is, for $i = 1, 2$ if $f, g \in L^{p_i}(\mathbb{R}^n)$ and $\lambda, \mu \in \mathbb{R}$ then

$$T(\lambda f + \mu g) = \lambda Tf + \mu Tg.$$

(this definition assumes that for $f \in L^{p_1} \cap L^{p_2}(\mathbb{R}^n)$ the Tf is well-defined).

Then if $p \in (p_1, p_2)$ the operator T naturally is defined for $f \in L^p(\mathbb{R}^n)$. Namely if we split $f = f_1 + f_2$ where $f_i \in L^{p_i}(\mathbb{R}^n)$ (we can always do that in view of Lemma 4.10) we set

$$Tf := Tf_1 + Tf_2.$$

Now assume that $f = \tilde{f}_1 + \tilde{f}_2$ is another decomposition but still $\tilde{f}_i \in L^{p_i}(\mathbb{R}^n)$, $i = 1, 2$. To make the operator well-defined we need to ensure that

$$T\tilde{f}_1 + T\tilde{f}_2 = Tf_1 + Tf_2 \quad a.e. \text{ in } \mathbb{R}^n. \tag{4.1}$$

Equivalently, we need to show

$$Tf_1 - T\tilde{f}_1 = T\tilde{f}_2 - Tf_2.$$

Observe that
$$f_1 - \tilde{f}_1 = f - f_2 - (f - \tilde{f}_2) = \tilde{f}_2 - f_2.$$
In particular, $f_1 - \tilde{f}_1 \in L^{p_1} \cap L^{p_2}$.

So, by linearity,
$$Tf_1 - T\tilde{f}_1 = T(f_1 - \tilde{f}_1) = T(\tilde{f}_2 - f_2) = T\tilde{f}_2 - Tf_2.$$
This establishes (4.1), that is Tf is well defined for each $f \in L^p(\mathbb{R}^n)$ as long as $p \in (p_1, p_2)$. There is more, we get also an estimate on $\|Tf\|_{L^p(\mathbb{R}^n)}$.

Theorem 4.11 (Marcienkiewicz Interpolation Theorem ("diagonal")). *Let $1 \leq p_1 < p_2 \leq \infty$*[18]*.*

Assume that T is a bounded linear operator from $L^{p_i}(\mathbb{R}^n)$ to $L^{p_i}(\mathbb{R}^n)$ for $i = 1, 2$. That is, for $i = 1, 2$

- *There exists $\Lambda_i > 0$ such that for any $f \in L^{p_i}(\mathbb{R}^n)$ we have $Tf \in L^{p_i}(\mathbb{R}^n)$ and moreover*
$$\|Tf\|_{L^{p_i}(\mathbb{R}^n)} \leq \Lambda_i \|f\|_{L^{p_i}(\mathbb{R}^n)}.$$
- *if $f, g \in L^{p_i}(\mathbb{R}^n)$ and $\lambda, \mu \in \mathbb{R}$ then*
$$T(\lambda f + \mu g) = \lambda Tf + \mu Tg.$$

Then for each $p \in (p_1, p_2)$ the operator T is a bounded linear operator from $L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, and we have
$$\|Tf\|_{L^p(\mathbb{R}^n)} \leq \Lambda_p \|f\|_{L^p(\mathbb{R}^n)}$$
where $\Lambda_q$is a constant depending only on Λ_1 and Λ_2 and θ.

Remark 4.12.
- The constant Λ_p here is not sharp in general, a technique called *complex interpolation* gives a sharper estimate
- We don't need linearity for T, only sublinearity, i.e. $Tf \leq Tf_1 + Tf_2$.
- Of course this works in much more generality.

Before we come to the proof of Theorem 4.11 we record a crucial tool, an amazing application of Fubini's theorem.

Proposition 4.13. *Fix $p \in [1, \infty)$ and $\Omega \subset \mathbb{R}^n$ be an open set.*

Let $f : \Omega \to \overline{\mathbb{R}}$ be $\mathcal{L}^n$-measurable. Then $f \in L^p(\Omega)$ if and only if
$$\lambda \mapsto \lambda^{p-1} \mathcal{L}^n (\{x \in \Omega : |f(x)| > \lambda\}) \quad \text{in } L^1((0, \infty)).$$
Moreover, in either case
$$\int_\Omega |f(x)|^p dx = p \int_{\lambda=0}^\infty \lambda^{p-1} \mathcal{L}^n (\{x \in \Omega : |f(x)| > \lambda\}) \, d\lambda.$$

[18] diagonal refers to the fact that $q_i = p_i$

Proof. Since $|f|$ is $\mathcal{L}^n$-measurable, also

$$\lambda \mapsto \lambda^{p-1}\mathcal{L}^n\left(\{x \in \Omega : |f(x)| > \lambda\}\right)$$

is $\mathcal{L}^1$-measurable (exercise).

So all we have to obtain is the identity.

$$\begin{aligned} &p\int_{\lambda=0}^{\infty} \lambda^{p-1}\mathcal{L}^n\left(\{x \in \Omega : |f(x)| > \lambda\}\right) d\lambda \\ =&p\int_{\lambda=0}^{\infty} \lambda^{p-1}\int_{\Omega} \chi_{\{|f(\cdot)|>\lambda\}}(x)dx\, d\lambda \end{aligned}$$

Now we use Fubini's theorem (rather: Tonelli), Theorem 4.4. One easily checks that $\lambda \times x \mapsto \lambda^{p-1}\chi_{\{|f(\cdot)|>\lambda\}}(x)$ is $\mathcal{L}^1 \times \mathcal{L}^n$-measurable (and nonnegative). So,

$$\begin{aligned} =&p\int_{\Omega}\int_{\lambda=0}^{\infty} \lambda^{p-1}\chi_{\{|f(\cdot)|>\lambda\}}(x)d\lambda\, dx \\ =&p\int_{\Omega}\int_{\lambda=0}^{|f(x)|} \lambda^{p-1}\, d\lambda\, dx \\ =&\int_{\Omega} \lambda^{p}\Big|_{\lambda=0}^{|f(x)|} dx \\ =&\int_{\Omega} |f(x)|^p\, dx. \end{aligned}$$

□

Proof of Theorem 4.11. Observe that $p \in (1, \infty)$.

Fix $f \in L^p(\mathbb{R}^n)$. For a fixed $\lambda > 0$ we apply Lemma 4.10 and split $f = f_1 + f_2$. Then $Tf = Tf_1 + Tf_2$ which is measurable.

Since $|Tf| \leq |Tf_1| + |Tf_2|$,

$$\{|Tf| > \lambda\} \subset \{|Tf_1| > \frac{\lambda}{2}\} \cup \{|Tf_2| > \frac{\lambda}{2}\}$$

so

$$\mathcal{L}^n\left(\{|Tf| > \lambda\}\right) \leq \mathcal{L}^n\left(\{|Tf_1| > \frac{\lambda}{2}\}\right) + \mathcal{L}^n\left(\{|Tf_2| > \frac{\lambda}{2}\}\right).$$

By Chebyshev inequality, Lemma 3.48, for $i = 1, 2$ and boundedness of T,

$$\begin{aligned} \mathcal{L}^n\left(\{|Tf_i| > \frac{\lambda}{2}\}\right) &\leq \left(\frac{\lambda}{2}\right)^{-p_i} \|Tf_i\|^{p_i}_{L^{p_i}(\mathbb{R}^n)} \\ &\leq \left(\frac{\lambda}{2}\right)^{-p_i} \|f_i\|^{p_i}_{L^{p_i}(\mathbb{R}^n)} \end{aligned}$$

By the definition of f_i we have

$$\left(\frac{\lambda}{2}\right)^{-p_1} \|f_1\|^{p_1}_{L^{p_1}(\mathbb{R}^n)} = \left(\frac{\lambda}{2}\right)^{-p_1} \int_{\{|f(\cdot)|>\lambda\}} |f(x)|^{p_1} dx$$

$$\left(\frac{\lambda}{2}\right)^{-p_2} \|f_2\|^{p_2}_{L^{p_2}(\mathbb{R}^n)} = \left(\frac{\lambda}{2}\right)^{-p_2} \int_{\{|f(\cdot)|\leq\lambda\}} |f(x)|^{p_2} dx$$

And thus

$$\mathcal{L}^n\left(\{|Tf| > \lambda\}\right) \leq \left(\frac{\lambda}{2}\right)^{-p_1} \int_{\{|f(\cdot)|>\lambda\}} |f(x)|^{p_1} dx + \left(\frac{\lambda}{2}\right)^{-p_2} \int_{\{|f(\cdot)|\leq\lambda\}} |f(x)|^{p_2} dx$$

This estimate holds for any $\lambda > 0$. In view of Proposition 4.13 we then have

$$\begin{aligned}
\|Tf\|^p_{L^p(\mathbb{R}^n)} \leq & p \int_{\lambda=0}^{\infty} \lambda^{p-1} \mathcal{L}^n\left(\{x \in \Omega : |Tf(x)| > \lambda\}\right) d\lambda \\
\leq & C \int_{\lambda=0}^{\infty} \lambda^{p-1} \left(\lambda^{-p_1} \int_{\{|f(\cdot)|>\lambda\}} |f(x)|^{p_1} dx + \lambda^{-p_2} \int_{\{|f(\cdot)|\leq\lambda\}} |f(x)|^{p_2} dx\right) d\lambda
\end{aligned}$$

Again by Fubini

$$\begin{aligned}
\|Tf\|^p_{L^p(\mathbb{R}^n)} \leq & C \int_{\mathbb{R}^n} |f(x)|^{p_1} \left(\int_{\lambda=0}^{|f(x)|} \lambda^{p-1} \lambda^{-p_1} d\lambda\right) dx \\
& + C \int_{\mathbb{R}^n} |f(x)|^{p_2} \left(\int_{\lambda=|f(x)|}^{\infty} \lambda^{p-1} \lambda^{-p_2} d\lambda\right) dx
\end{aligned}$$

Since $p_1 < p < p_2$ the integrals in λ converge and we have

$$\begin{aligned}
\|Tf\|^p_{L^p(\mathbb{R}^n)} \leq & C \int_{\mathbb{R}^n} |f(x)|^{p_1} \left(|f(x)|^p |f(x)|^{-p_1}\right) dx \\
& + C \int_{\mathbb{R}^n} |f(x)|^{p_2} \left(|f(x)|^p |f(x)|^{-p_2} d\lambda\right) dx \\
= & \tilde{C} \int_{\mathbb{R}^n} |f(x)|^p dx \\
= & \tilde{C} \|f\|^p_{L^p(\mathbb{R}^n)}.
\end{aligned}$$

We can conclude. □

Exercise 4.14. *Prove Theorem 4.11 under the weakened assumptions*

$$\|Tf\|_{L^{p_i}(\mathbb{R}^n)} \leq \Lambda_i \|f\|_{L^{p_i,\infty}(\mathbb{R}^n)}.$$

(Cf. Definition 3.49)

The ideas of the Marcienkiewicz-interpolation theorem can be vastly generalized. As particular version we record (without proof) the following "off-diagonal" version of Theorem 4.11. "off-diagonal" means that the L^p-spaces in the domain and target may not be the same. For a proof see [Grafakos, 2014, §1.4.4].

Theorem 4.15 (Marcienkiewicz Interpolation Theorem (off-diagonal))**.** *Let $1 \leq p_1 < p_2 \leq \infty$ and $1 \leq q_1 < q_2 \leq \infty$.*

Assume that T is a bounded linear operator from $L^{p_i}(\mathbb{R}^n)$ to $L^{q_i}(\mathbb{R}^n)$ for $i = 1, 2$. That is, for $i = 1, 2$

- *There exists $\Lambda_i > 0$ such that for any $f \in L^{p_i}(\mathbb{R}^n)$ we have $Tf \in L^{q_i}(\mathbb{R}^n)$ and moreover*
$$\|Tf\|_{L^{q_i}(\mathbb{R}^n)} \leq \Lambda_i \|f\|_{L^{p_i}(\mathbb{R}^n)}.$$
- *if $f, g \in L^{p_i}(\mathbb{R}^n)$ and $\lambda, \mu \in \mathbb{R}$ then*
$$T(\lambda f + \mu g) = \lambda Tf + \mu Tg.$$

Then for each $p \in (p_1, p_2)$, i.e. whenever $\theta \in (0,1)$ and
$$p = (1-\theta)p_1 + \theta p_2,$$
then for
$$q := (1-\theta)q_1 + \theta q_2$$
we that T is a bounded linear operator from $L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, and we have
$$\|Tf\|_{L^q(\mathbb{R}^n)} \leq \Lambda_q \|f\|_{L^p(\mathbb{R}^n)}$$
where Λ_q is a constant depending only on Λ_1 and Λ_2 and θ.

4.2. **Application: convolution.** Let us start with an observation that since the Lebesgue measure in invariant under translations and under the mapping $x \mapsto -x$ for any $f \in L^1(\mathbb{R}^n)$ and any $y \in \mathbb{R}^n$ we have
$$\int_{\mathbb{R}^n} u(x)dx = \int_{\mathbb{R}^n} u(x+y)dx = \int_{\mathbb{R}^n} u(-x)dx.$$
Also, since for every measurable set E and any $t > 0$ the set $tE = \{tE : x \in E\}$ has measure $\mathcal{L}^n(tE) = t^n \mathcal{L}^n(E)$, Theorem 1.81, we conclude that the Lebesgue integral has the following scaling property
$$\int_{\mathbb{R}^n} u(x/t)dx = t^n \int_{\mathbb{R}^n} u(x)dx \quad \forall t > 0.$$
Observe that in the case of the Riemann integral the above equalities are direct consequences of the change of variables formula. We will prove a corresponding change of variables formula for the Lebesgue integral later, but as for the proof of the above equalities we do not have to refer to the general change of variables formula as they follow directly from the properties of the Lebesgue measure mentioned above.

Definition 4.16. For measurable functions f and g on $\mathbb{R}^n$ we define the *convolution* by
$$f * g(x) := \int_{\mathbb{R}^n} f(x-y)\, g(y)\, dy$$

The convolution plays a crucial role in Analysis, in particular in Partial Differential equations,

- convolutions can be used to "mollify functions" (approximating non-differentiable functions by differentiable ones, we will discuss this below)

- representing solutions to linear differential equations, e.g. for $n \geq 3$
$$\Delta u(x) = f(x) \quad \text{in } \mathbb{R}^n$$
is (under suitable assumptions on f and u equivalent) to
$$u(x) = c \int_{\mathbb{R}^n} |x-y|^{2-n} f(y)\, dy.$$
(where c is a suitable constant)

To start with our analysis, the first question is under what conditions the convolution is well defined. If $f \in L^1(\mathbb{R}^n)$ and g is bounded, measurable and vanishes outside a bounded set, then the function $y \mapsto f(x-y)g(y)$ is integrable, so $(f*g)(x)$ is well defined and finite for every $x \in \mathbb{R}^n$. If $f, g \in L^1(\mathbb{R}^n)$, then it can happen that for a given x the function $y \mapsto f(x-y)g(y)$ is not integrable and hence $(f*g)(x)$ is not defined. However as a powerful application of the Fubini theorem we can prove the following surprising result.

Theorem 4.17. *If $f, g \in L^1(\mathbb{R}^n)$ then for a.e. $x \in \mathbb{R}^n$ the function $y \mapsto f(x-y)g(y)$ is integrable and hence $f*g(x)$ exists. Moreover $f*g \in L^1(\mathbb{R}^n)$ and*
$$\|f*g\|_{L^1(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.$$

Proof. The function $|f(x-y)g(y)|$ as a function of a variable $(x,y) \in \mathbb{R}^{2n}$ is measurable (because $(x,y) \mapsto f(x-y)$ and $(x,y) \mapsto g(y)$ are both $\mathcal{L}^2$-measurable). Hence Fubini's theorem (rather: Tonelli Theorem 4.4), implies
$$\begin{aligned}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)g(y)| dx\, dy &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)g(y)| dx\right) dy\\ &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)| dx\right) |g(y)| dy = \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.\end{aligned}$$
Therefore
$$\|f*g\|_{L^1(\mathbb{R}^n)} \leq \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)g(y)| dx\, dy \leq \|f\|_{L^1(\mathbb{R}^n)} \|g\|_{L^1(\mathbb{R}^n)}.$$
□

Theorem 4.17 is a special case of *Young's convolution inequality*

Theorem 4.18 (Young's convolution inequality)**.** *Let $p, q, r \in [1, \infty]$ satisfy the following relation*
$$1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$$
(where $\frac{1}{\infty} = 0$).

If $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$ then $f*g \in L^r(\mathbb{R}^n)$ and
$$\|f*g\|_{L^r(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}.$$

Exercise 4.19. *Proof the statement for $p = \infty$ or $q = \infty$ and $r = \infty$ (what does each imply for the other coefficients?).*

Proof. We consider only the case $1 < p, q, r < \infty$.

Set $s := \frac{p}{1-\frac{p}{r}}$, $t = \frac{q}{1-\frac{q}{r}}$. Observe that then

$$\frac{1}{r} + \frac{1}{s} + \frac{1}{t} = \frac{1}{r} + \left(\frac{1}{p} - \frac{1}{r}\right) + \left(\frac{1}{q} - \frac{1}{r}\right) = 1.$$

We write

$$|f(x-y)g(y)| = \underbrace{|f(x-y)|^{1-\frac{p}{r}}}_{\in L^s(dy)} \underbrace{|g(y)|^{1-\frac{q}{r}}}_{\in L^t(dy)} \underbrace{|f(x-y)|^{\frac{p}{r}}\, |g(y)|^{\frac{q}{r}}}_{L^r(dy)}.$$

From (generalized) Hölder's inequality we then have

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy \leq &\left(\int_{\mathbb{R}^n} |f(x-y)|^p dy\right)^{\frac{1-\frac{p}{r}}{p}} \left(\int_{\mathbb{R}^n} |g(y)|^q dy\right)^{\frac{1-\frac{q}{r}}{q}} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^q dy\right)^{\frac{1}{r}}\\ =&\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{1-\frac{q}{r}} (|f|^p * |g|^q)^{\frac{1}{r}}\end{aligned}$$

Now from the L^1-case,

$$\begin{aligned}\| (|f|^p * |g|^q)^{\frac{1}{r}} \|_{L^r(\mathbb{R}^n)} =&\| |f|^p * |g|^q\|_{L^1(\mathbb{R}^n)}^{\frac{1}{r}} \leq \left(\| |f|^p\|_{L^1(\mathbb{R}^n)}\, \||g|^q\|_{L^1(\mathbb{R}^n)}\right)^{\frac{1}{r}}\\ =&\|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{\frac{q}{r}}\end{aligned}$$

Plugging all this together we find

$$\begin{aligned}\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy\right)^r dx \leq &\left(\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{1-\frac{q}{r}} \| (|f|^p * |g|^q)^{\frac{1}{r}} \|_{L^r(\mathbb{R}^n)}\right)^r\\ \leq &\left(\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{r-\frac{q}{r}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{\frac{q}{r}}\right)^r\\ = &\left(\|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}\right)^r\end{aligned}$$

Thus,

$$\|f * g\|_{L^r(\mathbb{R}^n)}^r \leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy\right)^r dx \leq \left(\|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}\right)^r.$$

Taking this inequality to the power $()^{\frac{1}{r}}$ we conclude. $\square$

Now let us look at algebraic properties of the convolution

Exercise 4.20. $(L^1(\mathbb{R}^n), *)$ *is a commutative algebra. Namely, if* $\alpha, \beta \in \mathbb{R}$, $f, g, h \in L^1(\mathbb{R}^n)$ *then*

(1) $f * g = g * f$
(2) $f * (g * h) = (f * g) * h$
(3) $f * (\alpha g + \beta h) = \alpha f * g + \beta f * h.$

Hint: *Fubini*

Let us also record

Exercise 4.21. *Let $f, g \in L^1(\mathbb{R}^n)$ and $h \in C_c^\infty(\mathbb{R}^n)$ then*
$$\int_{\mathbb{R}^n} (f * g)\ h = \int_{\mathbb{R}^n} f\ (\bar{g} * h)$$
where
$$\bar{g}(x) := g(-x).$$

Hint: *Fubini*

So, nice algebraic structure, nice! But now lets move towards serious business, namely "mollification".

For the "mollification aspect" we need a further definition

Definition 4.22.
(1) We say that $f \in L^p_{loc}(\mathbb{R}^n)$ if for any $x \in \mathbb{R}^n$ there exists some ball $B(x,r)$ such that $f \in L^p(B(x,r))$.
(2) For $g \in C^0(\mathbb{R}^n)$ the *support* is "where the function lives", more precisely
$$\operatorname{supp} g := \overline{\{x \in \mathbb{R}^n : g(x) \neq 0\}}.$$
A continuous function g is said to have *compact support* if $\operatorname{supp} g$ is compact.
(3) This can be generalized for $\mathcal{L}^n$-measurable $g : \mathbb{R}^n \to \overline{\mathbb{R}}$.
Fix such a g, and define the family of (measurable!) $\mathcal{S}$ as follows
$$S \in \mathcal{S} \quad :\Leftrightarrow S \text{ is closed and } g(x) = 0 \text{ for } \mathcal{L}^n\text{-a.e. } x \in \mathbb{R}^n \setminus S$$
$$\operatorname{supp} g := \bigcap_{S \in \mathcal{S}} S.$$
So a $\mathcal{L}^n$-measurable function is said to have *compact support* if and only if $\operatorname{supp} g$ is compact.

Exercise 4.23. *Show that whenever $f \in L^p_{loc}(\mathbb{R}^n)$ then $f \in L^q_{loc}(\mathbb{R}^n)$ for any $1 \leq q \leq p$*

Hint: *Hölder's inequality.*

Exercise 4.24. *Show that the two definitions of support in Definition 4.22 coincide for continuous functions.*

If $f \in L^1_{loc}(\mathbb{R}^n)$ and g is bounded, measurable and vanishes outside a bounded set, then the function $y \mapsto f(x-y)g(y)$ is integrable, so $(f * g)(x)$ is well defined and finite for every $x \in \mathbb{R}^n$. If g is better, then $f * g$ is as good as g.

Theorem 4.25. *If $f \in L^1_{loc}(\mathbb{R}^n)$ and g is continuous with compact support then*

(1) $f * g$ is continuous on $\mathbb{R}^n$
(2) If additionally $g \in C^k(\mathbb{R}^n)$ then $f * g \in C^k(\mathbb{R}^n)$, $k \in \mathbb{N}$, and we have
$$\partial_\alpha (f * g)(x) = (f * \partial_\alpha g)(x) \quad \alpha \text{ multiindex: } |\alpha| \leq k.$$

Proof. (1) Since g is bounded and has compact support, $(f * g)(x)$ is well defined and finite for all $x \in \mathbb{R}^n$. Fix $x \in \mathbb{R}^n$. The function $y \mapsto f(y)g(x-y)$ vanishes outside a sufficiently large ball B (because g as compact support). Let $x_n \to x$. Then there is a ball B (perhaps larger than the one above), so that all the functions

$$y \mapsto f(y)g(x_n - y)$$

vanish outside B. Hence

$$|f(y)g(x_n - y)| \leq \|g\|_{L^\infty} |f(y)| \chi_B(y) \in L^1(\mathbb{R}^n)$$

and the dominated convergence theorem, Theorem 3.26, yields

$$(f * g)(x_n) = \int_{\mathbb{R}^n} f(y)g(x_n - y)dy \to \int_{\mathbb{R}^n} f(y)g(x-y)dy = (f * g)(x)$$

which proves continuity of $f * g$.

(2) Pick $i \in \{1, \ldots, n\}$ and $|h| \leq 1$ and fix $x \in \mathbb{R}^n$

We have (observe all terms exists since $g \in C^1(\mathbb{R}^n)$ and $\partial_i g$ still has compact support)

$$\begin{aligned}&(f * g)(x + he_i) - (f * g)(x) - (f * \partial_i g)(x)h \\ =&\int_{\mathbb{R}^n} f(y)\left(g(x + he_i - y) - g(x-y) - \partial_i g(x-y)h\right) dy\end{aligned}$$

Observe that if $\operatorname{supp} g$ is compact then $\operatorname{supp} g(z - \cdot)$ is still compact (for a fixed z). It is easy to check that the following set is compact

$$K_x := \bigcup_{|z| \leq 1} \operatorname{supp} g(x + z - \cdot).$$

Moreover,

$$g(x + he_i - y) - g(x - y) - \partial_i g(x-y)h = 0 \quad y \notin K_x.$$

Since g is continuously differentiable on $\mathbb{R}^n$ we have

$$|g(x + he_i - y) - g(x-y) - \partial_i g(x-y)h| \leq C_x |h| \quad \forall y \in \mathbb{R}^n, |h| \leq 1.$$

and moreover by differentiability for every y

$$\frac{|g(x + he_i - y) - g(x-y) - \partial_i g(x-y)h|}{|h|} \xrightarrow{|h| \to 0} 0.$$

Thus

$$\begin{aligned}&\frac{(f * g)(x + he_i) - (f * g)(x) - (f * \partial_i g)(x)h}{|h|} \\ =&\int_{\mathbb{R}^n} f(y) \frac{g(x + he_i - y) - g(x-y) - \partial_i g(x-y)h}{|h|} dy\end{aligned}$$

and

$$f(y) \frac{g(x + he_i - y) - g(x-y) - \partial_i g(x-y)h}{|h|} \xrightarrow{|h| \to 0} 0 \text{pointwise for a.e. } y$$

and
$$|f(y)\frac{g(x+he_i-y)-g(x-y)-\partial_i g(x-y)h}{|h|}| \leq C|f(y)|$$
We conclude by the dominated convergence theorem that
$$\frac{(f*g)(x+he_i)-(f*g)(x)-(f*\partial_i g)(x)h}{|h|} \xrightarrow{|h|\to 0} 0.$$
Thus $f*g$ has a partial derivative at every x. That by itself is not enough to conclude differentiability! However, since morevoer $f*\partial_i g$ is continuous (by (1)) we can conclude that $f*g$ is continuously differentiable everywhere. Repeating this argument for higher derivatives gives the claim.

$\square$

Let $\eta \in C_c^\infty(B(0,1))$ be a function such that $\eta \geq 0$ and $\int_{\mathbb{R}^n} \eta(x) = 1$. E.g. take

$$\tilde{\eta}(x) := \begin{cases} e^{\frac{1}{|x|^2-1}} & |x| \leq 1 \\ 0 & |x| \geq 1 \end{cases} \tag{4.2}$$

One can show that $\tilde{\eta} \in C_c^\infty(B(0,1))$, and we set $\eta := \left(\int_{\mathbb{R}^n} \tilde{\eta}\right)^{-1} \tilde{\eta}$. η is often called a *bump function* (because thats what it is) or a *mollifier* (reason: below).

For $\varepsilon > 0$ we set
$$\eta_\varepsilon(x) := \varepsilon^{-n}\eta(x/\varepsilon)$$

Exercise 4.26. *Show that* $\operatorname{supp} \eta_\varepsilon \subset B(0,\varepsilon)$, $\eta_\varepsilon \geq 0$ *and* $\int_{\mathbb{R}^n} \eta_\varepsilon = 1$.

For $f \in L^1_{loc}(\mathbb{R}^n)$ we set
$$f_\varepsilon := f * \eta_\varepsilon.$$
In view of Theorem 4.25, $f_\varepsilon \in C^\infty(\mathbb{R}^n)$. This is called the *mollification* of f.

Observe that as $\varepsilon \to 0$ we have in a very handwaving sense
$$\eta_\varepsilon(x) \xrightarrow{\varepsilon\to 0} \begin{cases} 0 & x \neq 0 \\ \infty \cdot \text{“}\eta(0)\text{''} & x = 0 \end{cases}$$
More precisely we have for any measurable $A \subset \mathbb{R}^n$
$$\mathcal{L}^n \llcorner \eta_\varepsilon \xrightarrow{*} \eta(0)\delta_0$$
where δ_0 is the *Dirac measure*
$$\delta_0(A) = \begin{cases} 1 & 0 \in A \\ 0 & \text{otherwise.} \end{cases}$$
Above, (weak*) convergence is understood in the following sense:

Lemma 4.27. *For any $f \in C^0(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$ we have*

$$\int_{\mathbb{R}^n} f d\left(\mathcal{L}^n \llcorner \eta_\varepsilon\right) \xrightarrow{\varepsilon \to 0} \int_{\mathbb{R}^n} f d\delta_0 = f(0).$$

Proof. We have

$$\int_{\mathbb{R}^n} f d\left(\mathcal{L}^n \llcorner \eta_\varepsilon\right) = \int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)\, d\mathcal{L}^n.$$

So all we need to show is that

$$\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)\, d\mathcal{L}^n \xrightarrow{\varepsilon \to 0} f(0).$$

Now observe since $\int \eta_\varepsilon = 1$ we have

$$\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)dx - f(0) = \int_{\mathbb{R}^n} \left(f(x) - f(0)\right)\eta_\varepsilon(x)dx.$$

Since f is continuous at 0, for any $\delta > 0$ there exists a radius $r > 0$ such that $|f(x) - f(0)| \leq \delta$ for all $x \in B(0,r)$. If we take $\varepsilon < r$ we then have

$$\left|\int_{\mathbb{R}^n} \left(f(x) - f(0)\right)\eta_\varepsilon(x)dx\right| \leq \delta \int_{\mathbb{R}^n} \eta_\varepsilon(x)dx = \delta.$$

This holds for any δ and $\varepsilon < r(\delta)$ so we have

$$\lim_{\varepsilon \to 0}\left|\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(x)dx - f(0)\right| = 0.$$

□

What does this mean for $f_\varepsilon(x) = f * \eta_\varepsilon(x) = \int f(x-y)\eta_\varepsilon(y)$? It converges to $f(x)$. And it does so quite nicely.

Theorem 4.28. *Let $f : \mathbb{R}^n \to \mathbb{R}$.*

(1) If f is continuous then f_ε converges uniformly to f on compact sets as $\varepsilon \to 0$, i.e. for all K compact

$$\|f - f_\varepsilon\|_{L^\infty(K)} \xrightarrow{\varepsilon \to 0} 0.$$

(2) If $f \in C^k(\mathbb{R}^n)$ then $D^\alpha f_\varepsilon$ converges uniformly to $D^\alpha f$ on compact sets for any $|\alpha| \leq r$ as $\varepsilon \to 0$, i.e. for all K compact

$$\max_{|\alpha| \leq k} \|D^\alpha f - D^\alpha f_\varepsilon\|_{L^\infty(K)} \xrightarrow{\varepsilon \to 0} 0.$$

Proof. (1) Uniform continuity of f on compact sets implies that for any compact set K

$$\sup_{|y|<\varepsilon} \sup_{x \in K} |f(x) - f(x-y)| \to 0 \quad \text{as } \varepsilon \to 0.$$

Since $\int_{\mathbb{R}^n} \eta_\varepsilon(y)dy = 1$ we have $\int_{\mathbb{R}^n} f(x)\eta_\varepsilon(y)dy = f(x)$ and hence for any compact set K,

$$\begin{aligned}\sup_{x\in K} |f(x) - f_\varepsilon(x)| =& \sup_{x\in K} \left|\int_{\mathbb{R}^n} (f(x) - f(x-y))\eta_\varepsilon(y)dy\right| \\ \le& \sup_{x\in K} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|\, \eta_\varepsilon(y)dy \to 0 \quad \text{as } \varepsilon \to 0.\end{aligned}$$

(2) Follows from (1) by an induction argument and the fact tat $D^\alpha(f*\eta_\varepsilon) = (D^\alpha f)*\eta_\varepsilon$.

□

Exercise 4.29. *(1) If $f \in C^\alpha$ for $\alpha \in (0,1]$ show that f_ε converges to f in C^β for any $\beta < \alpha$, i.e.*

$$\sup_{x\neq y} \frac{|f(x) - f_\varepsilon(x) - f(y) - f_\varepsilon(y)|}{|x-y|^\beta} \xrightarrow{\varepsilon\to 0} 0.$$

(2) Show that this may not be true for $\beta = \alpha = 1$ (it is also not necessarily true for $\beta = \alpha < 1$)

Hint: Lipschitz functions may not be everywhere differentiable

Now we want to get convergence also in $L^p(\mathbb{R}^n)$.

Lemma 4.30. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p \le \infty$ then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f_\varepsilon\|_{L^p(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)}$.*

Proof. Follows from Theorem 4.18:

$$\|f_\varepsilon\|_{L^p(\mathbb{R}^n)} = \|f*\eta_\varepsilon\|_{L^p(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)} \underbrace{\|\eta_\varepsilon\|_{L^1(\mathbb{R}^n)}}_{=1}.$$

□

Theorem 4.31. *If $f \in L^p(\mathbb{R}^n)$ and $1 \le p<\infty$ then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f - f_\varepsilon\|_{L^p(\mathbb{R}^n)} \xrightarrow{\varepsilon\to 0} 0$.*

For the proof of Theorem 4.31 we need the following lemma

Lemma 4.32. *If $f \in L^p(\mathbb{R}^n)$, $1 \le p < \infty$ then*

$$\lim_{y\to 0} \int_{\mathbb{R}^n} |f(x+y) - f(x)|^p dx = 0.$$

Proof. For $y \in \mathbb{R}^n$ let $\tau_y : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ be the translation operator defined by

$$(\tau_y f)(x) = f(x+y) \quad \text{for } f \in L^p(\mathbb{R}^n).$$

The lemma claims that τ_y is continuous, i.e. $\|\tau_y f - f\|_{L^p} \xrightarrow{y\to 0} 0$ as $y \to 0$. Given $\varepsilon > 0$ let g be a compactly supported continuous function such that $\|f-g\|_{L^p} < \varepsilon/3$, see Theorem 3.28. Then

$$\begin{aligned}\|\tau_y f - f\|_{L^p} \le& \|\tau_y f - \tau_y g\|_{L^p} + \|\tau_y g - g\|_{L^p} + \|f-g\|_{L^p} = 2\|f-g\|_{L^p} + \|\tau_y g - g\|_{L^p} \\ <& 2\varepsilon/3 + \|\tau_y g - g\|_{L^p}\end{aligned}$$

Since $\tau_y g \to g$ converges uniformly, and $K := \bigcup_{|y|\leq 1} \operatorname{supp} \tau_y g$ is bounded, there is $\delta > 0$ such that

$$\|\tau_y g - g\|_{L^p(\mathbb{R}^n)} = \|\tau_y g - g\|_{L^p(K)} \overset{\text{Hölder}}{\leq} C(K)\,\|\tau_y g - g\|_{L^\infty(\mathbb{R}^n)} < \varepsilon/3 \quad \text{for } |y| < \delta.$$

Hence,

$$\|\tau_y f - f\|_{L^p} < \frac{\varepsilon}{3} \quad \text{for } |y| < \delta,$$

which proves the lemma. □

Proof of Theorem 4.31. Assume first that $1 < p < \infty$. Hölder's inequality and Fubini's theorem yield

$$\begin{aligned}
\|f - f_\varepsilon\|_{L^p} &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\eta_\varepsilon(y)dy \right|^p dx \\
&\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x) - f(x-y)| \eta_\varepsilon(y)^{1/p} \eta_\varepsilon(y)^{1/q} dy \right)^p dx \\
&\leq \int_{\mathbb{R}^n} \left(|f(x) - f(x-y)|^p \eta_\varepsilon(y) dy \right) \underbrace{\left(\int_{\mathbb{R}^n} \eta_\varepsilon(y) dy \right)^{\frac{p}{q}}}_{=1} dx \\
&= \int_{\mathbb{R}^n} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|^p \eta_\varepsilon(y) dy dx \\
&= \int_{B(0,\varepsilon)} \left(\int_{\mathbb{R}^n} |f(x) - f(x-y)|^p dx \right) \eta_\varepsilon(y)\, dy \xrightarrow{\varepsilon \to 0} 0.
\end{aligned}$$

In the last step we used Lemma 4.32. If $p = 1$, then the above argument simplifies, because we do not have to use Hölder's inequality. □

Corollary 4.33. *Let $1 \leq p < \infty$. If $f \in L^p(\Omega)$ for an open set Ω then for any open compactly contained $\Omega' \subset\subset \Omega$ (i.e. Ω' is open, $\overline{\Omega'}$ is compact and $\overline{\Omega'} \subset \Omega$) we have $f_\varepsilon \in C^\infty(\overline{\Omega'})$ for all small $\varepsilon \ll 1$ and $\|f_\varepsilon - f\|_{L^p(\Omega')} \xrightarrow{\varepsilon \to 0} 0$.*

Proof. Set $d := \frac{\operatorname{dist}(\Omega', \partial\Omega)}{100} > 0$. Let $\bar\eta \in C_c^\infty(B(\Omega', 50d), [0,\infty))$ be a bump function with $\bar\eta \equiv 1$ in $B(\Omega', 20d)$[19]

Set

$$\bar f := \bar\eta f.$$

Clearly $\bar f \in L^p(\mathbb{R}^n)$, and by Theorem 4.31 we have that $\bar f \in C^\infty(\mathbb{R}^n)$ and

$$\|\bar f_\varepsilon - \bar f\|_{L^p(\mathbb{R}^n)} \xrightarrow{\varepsilon \to 0} = 0.$$

In particular $\bar f_\varepsilon \in C^\infty(\overline{\Omega'})$ and (using that $\bar\eta \equiv 1$ in Ω')

$$\|\bar f_\varepsilon - f\|_{L^p(\Omega')} = \|\bar f_\varepsilon - \bar f\|_{L^p(\Omega')} \xrightarrow{\varepsilon \to 0} = 0.$$

[19]Here: $B(A,r) := \{x \in \mathbb{R}^n : \operatorname{dist}(x, A) < r\}$

We now claim that for $\varepsilon < d$ we have $\bar{f}_\varepsilon = f_\varepsilon$ for $x \in \Omega'$. Indeed if $x \in \Omega'$ then denoting the mollification kernel for f_ε by $\eta \in C_c^\infty(B(0,1))$, $f_\varepsilon = f * \varepsilon^{-n}\eta(\cdot/\varepsilon)$ then $\operatorname{supp} \varepsilon^{-n}\eta((x-\cdot)/\varepsilon) \subset B(x,\varepsilon) \subset B(\Omega', d)$. Consequently, for any $x \in \Omega'$

$$\begin{aligned} f_\varepsilon(x) =& \varepsilon^{-n} \int_{\mathbb{R}^n} f(y)\eta((x-y)/\varepsilon)dy = \varepsilon^{-n} \int_{B(\Omega',d)} f(y)\eta((x-y)/\varepsilon)dy \\ =& \varepsilon^{-n} \int_{B(\Omega',d)} \bar{\eta}(y) f(y)\eta((x-y)/\varepsilon)dy = \bar{f}_\varepsilon(x). \end{aligned}$$

Thus

$$\|f_\varepsilon - f\|_{L^p(\Omega')} = \|\bar{f}_\varepsilon - f\|_{L^p(\Omega')} \xrightarrow{\varepsilon \to 0} = 0.$$

The proof is complete. $\square$

One can use convolution to obtain another proof of density, see Theorem 3.32(3). See Lemma 4.38 for a version for any open set $\Omega \subset \mathbb{R}^n$.

Exercise 4.34. *Use convolution to give another proof of Theorem 3.32(3) on $\mathbb{R}^n$:*

Namely show that $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for all $1 \leq p < \infty$. That is for any $f \in L^p(\mathbb{R}^n)$ there exists $(g_k)_{k \in \mathbb{N}} \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|g_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k \to \infty} 0.$$

(1) Approximate f by convolution with some $f_\varepsilon \in C^\infty(\mathbb{R}^n)$ (but there is no reason that $f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$)

(2) Take $\bar{\eta} \in C_c^\infty(B(0,2))$, $\bar{\eta} \equiv 1$ in $B(0,1)$ and $\|\bar{\eta}\|_{L^\infty(\mathbb{R}^n)} \leq 1$ a bump function and set

$$g_k(x) := \bar{\eta}(x/k)\, f_{\frac{1}{k}}(x).$$

(3) Show that

$$\|\bar{\eta}(x/k)\left(f_{1/k} - f\right)\|_{L^p(\mathbb{R}^n)} \xrightarrow{k \to \infty} 0.$$

(4) Show (Hint: dominated convergence!)

$$\|(1 - \bar{\eta}(x/k))f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k \to \infty} 0.$$

(5) Conclude that

$$\|g_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k \to \infty} 0.$$

Exercise 4.35. *Show that Exercise 4.34 is false for $p = \infty$*

(Hint: continuous functions)

Exercise 4.36. *Let $p \in [1, \infty)$. Show that for any $f \in L^p(\mathbb{R}^n)$ with $f \geq 0$ a.e. there exists $f_k \in C_c^\infty(\mathbb{R}^n)$, $f_k \geq 0$ everywhere such that*

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k \to \infty} 0.$$

Proposition 4.37. *Let $p \in [1, \infty)$ and $f : \mathbb{R}^n \to \mathbb{R}$ be any function. Then the following are equivalent:*

(1) $f \in L^p(\mathbb{R}^n)$
(2) There exists $f_k \in C_c^\infty(\mathbb{R}^n)$ *with* $f_k \to f$ *a.e. and*

$$\sup_k \|f_k\|_{L^p(\mathbb{R}^n)} < \infty.$$

In either case we have

$$\|f\|_{L^p(\mathbb{R}^n)} = \inf_{(f_k)_k} \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} \tag{4.3}$$

where the infimum is taken over all sequences $(f_k)_k$ *as in* (2)*.*

Proof. $\underline{\Rightarrow}$ This follows from Theorem 4.31. Since $f \in L^p(\mathbb{R}^n)$ we find approximations $f_k \in C_c^\infty(\mathbb{R}^n)$ and

$$\|f_k\|_{L^p(\mathbb{R}^n)} \leq \|f - f_k\|_{L^p(\mathbb{R}^n)} + \|f\|_{L^p(\mathbb{R}^n)}.$$

We also get one part of the inequality

$$\inf_{(\tilde{f}_k)_k} \liminf_{k\to\infty} \|\tilde{f}_k\|_{L^p(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|f-f_k\|_{L^p(\mathbb{R}^n)}+\liminf_{k\to\infty} \|f\|_{L^p(\mathbb{R}^n)} = 0+\|f\|_{L^p(\mathbb{R}^n)}.$$

$\underline{\Leftarrow}$ Let f_k be as in (2). In particular f_k are $\mathcal{L}^n$-measurable, so as a pointwise limit f is $\mathcal{L}^n$-measurable, Theorem 2.13. In particular $|f|$ is $\mathcal{L}^n$-measurable.

We can apply Fatou's lemma, Corollary 3.9, to $|f_k|^p$ and by pointwise convergence

$$\int_{\mathbb{R}^n} |f|^p = \int_{\mathbb{R}^n} \lim_{k\to\infty} |f_k|^p \leq \liminf_{k\to\infty} \int_{\mathbb{R}^n} |f_k|^p.$$

Since $\int_{\mathbb{R}^n} |f|^p < \infty$ we have $f \in L^p(\mathbb{R}^n)$. The above inequality holds for any approximating sequence $(f_k)_k$ so we get the remaining part of (4.3), namely

$$\|f\|_{L^p(\mathbb{R}^n)} \leq \inf_{(f_k)_k} \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)}$$

□

One can also get another proof of Theorem 3.32(3) on any Ω.

Lemma 4.38. *Let* $\Omega \subset \mathbb{R}^n$ *open and* $f \in L^p(\Omega)$*. Then there exist* $f_k \in C_c^\infty(\Omega)$ *such that*

$$\|f_k - f\|_{L^p(\Omega)} \xrightarrow{k\to\infty} 0.$$

Proof. Extend f by zero outside of Ω. We first show that there exist $g_k \in L^p(\mathbb{R}^n)$ with $\operatorname{supp} g_k \subset \Omega$ and

$$\|f - g_k\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Any open set can be decomposed into a countable union of closed dyadic cubes $\bigcup_{i=1}^\infty Q_i$ with pairwise disjoint interior. Let $g_k := \chi_{\bigcup_{i=1}^k Q_i} f$. By the dominated convergence theorem $\|g_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$. Observe that $\operatorname{supp} g_k \subset \bigcup_{i=1}^k Q_i$ which is a closed subset inside Ω. Since Ω is open $\delta := \operatorname{dist}(\operatorname{supp} g_k, \mathbb{R}^n\backslash\Omega) > 0$.

Now set $f_{k,\sigma} := \eta_\sigma * g_k$, for $\sigma < \delta$. Fix $\varepsilon > 0$. Then there exists $\sigma > 0$ so small such that

$$\|\eta_\sigma * f - f\|_{L^p(\mathbb{R}^n)} \leq \frac{\varepsilon}{2}.$$

Also we can take $k \in \mathbb{N}$ (independent of σ actually) such that

$$\|\eta_\sigma * g_k - \eta_\sigma * f\| \leq \|g_k - f\| < \frac{\varepsilon}{2}.$$

Then we have $f_{k,\sigma} \in C_c^\infty(\Omega)$ and

$$\|f_{k,\sigma} - f\|_{L^p(\mathbb{R}^n)} \leq \|\eta_\sigma * g_k - \eta_\sigma * f\| + \|\eta_\sigma * f - f\| < \varepsilon.$$

□

Exercise 4.39 (Censored Mollification). *Take three radii $0 < r < \rho < R$ and assume $u \in C^0(\overline{B(0,R)})$. Show that there is an approximation $u_k \in C^0(\overline{B(0,R)})$ with*

$$\|u_k - u\|_{L^\infty(B(0,R))} \xrightarrow{k\to\infty} 0$$

that satisfies the following conditions for all $k \in \mathbb{N}$

- $u_k \equiv u$ *in* $B(0,R) \setminus B(0,\rho)$
- $u_k \in C^\infty(B(0,r))$

For this use the following definition of a "censored" mollification

$$u_\delta(x) := \int_{\mathbb{R}^n} \eta(z) u(x + \delta\theta(x)z) dz,$$

for some choice of $\theta \in C_c^\infty(B(0,\rho),[0,1])$ and $\theta \equiv 1$ in $B(0,r)$, and a typical bump function $\eta \in C_c^\infty(B(0,2))$, $\eta \equiv 1$ in $B(0,1)$ and $\int \eta = 1$.

4.3. **A first glimpse on Sobolev spaces.** From the approximation Exercise 4.34 we have the following equivalent definition of $L^p(\mathbb{R}^n)$.

Proposition 4.40. *Let $p \in [1,\infty)$.*

Consider the space $\mathcal{L}^p(\mathbb{R}^n)$ as all $f \in C_c^\infty(\mathbb{R}^n)$ with the norm $\|f\|_{L^p(\mathbb{R}^n)}$. Every norm induces a metric $d_{\mathcal{L}^p}(f,g) := \|f - g\|_{L^p(\mathbb{R}^n)}$. Denote by $\tilde{L}^p(\mathbb{R}^n)$ the metric completion of $\mathcal{L}^p(\mathbb{R}^n)$ under the metric $d_{\mathcal{L}^p}(f,g)$.

Then $\tilde{L}^p(\mathbb{R}^n) = L^p(\mathbb{R}^n)$ in the following sense:

The uniformly continuous linear functional $\mathrm{id} : \mathcal{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ *given by* $\mathrm{id}\, f := f$ *extends to a (unique) isometric isomorphism $\tilde{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$.*

Proof. Since (by definition) $\mathcal{L}^p(\mathbb{R}^n)$ is dense in $\tilde{L}^p(\mathbb{R}^n)$ and id is clearly Lipschitz continuous with Lipschitz constant 1 (and thus uniformly continuous) id extends uniquely to a continuous map $\mathrm{id} : \tilde{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, given by $\mathrm{id}\, f := \lim_{k\to\infty} \mathrm{id}\, f_k$ where f_k is any sequence in $\mathcal{L}^p(\mathbb{R}^n)$ converging to f.

Clearly id is injective and isometric. It only needs to be shown that it is onto. So let $f \in L^p(\mathbb{R}^n)$ then by Exercise 4.34 there exists $f_k \in L^p(\mathbb{R}^n)$ with $\|f_k - f\|_{L^p} \to 0$. But then f_k is a Cauchy sequence also in $\mathcal{L}^p$, and thus $\mathrm{id}\, f_k = f_k$ converges on both sides. □

Exercise 4.41. *Let $p \in [1, \infty]$. Show that $(C_c^\infty(\mathbb{R}^n), \|\cdot\|_{W^{1,p}(\mathbb{R}^n)})$ is a normed space where*

$$\|f\|_{W^{1,p}(\mathbb{R}^n)} := \|f\|_{L^p(\mathbb{R}^n)} + \|\nabla f\|_{L^p(\mathbb{R}^n)}.$$

This leads to a (first, there are several) definition of *Sobolev spaces* $W^{1,p}(\mathbb{R}^n)$.

Definition 4.42. Let $p \in [1, \infty)$. Consider $C_c^\infty(\mathbb{R}^n)$ equipped with the norm

$$\|f\|_{W^{1,p}(\mathbb{R}^n)} := \|f\|_{L^p(\mathbb{R}^n)} + \|\nabla f\|_{L^p(\mathbb{R}^n)}.$$

We denote $W^{1,p}(\mathbb{R}^n)$ as the metric completion of $C_c^\infty(\mathbb{R}^n)$ under this norm.

So what kind of functions are in this Sobolev space $W^{1,p}(\mathbb{R}^n)$? Since the $\|f\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{W^{1,p}(\mathbb{R}^n)}$ we see that $W^{1,p}(\mathbb{R}^n)$ must be a subspace of $L^p(\mathbb{R}^n)$, i.e. a first equivalent notion is

Definition 4.43. Let $p \in [1, \infty)$. $f \in W^{1,p}(\mathbb{R}^n)$ if and only if $f \in L^p(\mathbb{R}^n)$ and there exists $f_k \in C_c^\infty(\mathbb{R}^n)$ such that f_k is a $W^{1,p}(\mathbb{R}^n)$-Cauchy sequence and

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Alright, so what does that mean: $W^{1,p}$ consists of L^p-functions that satisfy an additional condition: their *distributional derivative* belongs also to L^p. That needs some explanation.

Definition 4.44. A *distribution* is a linear map $T : C_c^\infty(\mathbb{R}^n) \to \mathbb{R}$ (it should also be continuous, but we discuss this later, when we talk about distributions in detail).

The (distributional) derivative of a distribution T is given by

$$\partial_\alpha T(\varphi) := (-1)^\alpha T(\partial_\alpha \varphi), \quad \varphi \in C_c^\infty(\mathbb{R}^n).$$

Every L^1_{loc}-function f is (equivalent) to a distribution:

Usually we think of functions $f : \mathbb{R}^n \to \mathbb{R}$ as function evaluated at some point x.

If we know for all $x \in \mathbb{R}^n$ what $f(x)$ looks like, then we know what f is. More precisely, *we can test when to functions are the same: $f = g$ if $f(x) = g(x)$ for all $x \in \mathbb{R}^n$*

But we can evaluate functions differently, with other test-functions. For example, for $\varphi \in C_c^\infty(\mathbb{R}^n)$, we can consider f as *distribution*, i.e.

$$f[\varphi] := \langle f, \varphi \rangle := \int f(x)\varphi(x)dx$$

Still: *if we know for all $\varphi \in C_c^\infty(\mathbb{R}^n)$ what $\langle f, \varphi \rangle$ looks like, then we know what f is.*

More precisely *we can test when to functions are the same: $f = g$ if $\langle f, \varphi \rangle = \langle g, \varphi \rangle$ for all $\varphi \in C_c^\infty(\mathbb{R}^n)$*

Lemma 4.45 (Fundamental lemma of the Calculus of Variations)**.** *Let $f, g \in L^1_{loc}(\mathbb{R}^n)$ and assume that*

$$\langle f, \varphi \rangle = \langle g, \varphi \rangle \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$$

then $f = g$ almost everywhere.

In particular if

$$\langle f, \varphi \rangle = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n),$$

then $f = 0$ almost everywhere.

Proof. Since

$$\langle f, \varphi \rangle = \langle g, \varphi \rangle \quad \Leftrightarrow \quad \langle f - g, \varphi \rangle = 0,$$

we only need to show that if

$$\langle f, \varphi \rangle = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n),$$

then $f = 0$ a.e.

First assume that f is continuous around a point $x_0 \in \mathbb{R}^n$. If $f(x_0) \neq 0$ then w.l.o.g. $f(x_0) > 0$. By continuity there exists a small ball $B(x_0, r)$ such that $f(x) \geq \frac{1}{2} f(x_0)$ for all $x \in B(x_0, r)$.

Now let φ be a bump function in $B(x_0, r)$, i.e. $\varphi \in C_c^\infty(B(x_0, r))$, $\varphi \geq 0$ everywhere and $\inf_{B(x_0, r/2)} \varphi > 0$; e.g. $\varphi = \tilde{\varphi}((x - x_0)/r)$ where $\tilde{\varphi}$ is from (4.2).

Then

$$0 = \langle f, \varphi \rangle = \int_{B(x_0,r)} f\varphi \geq \frac{1}{2} f(x_0) \int_{B(x_0,r)} \varphi > 0,$$

a contradiction. So $f(x_0) = 0$. This holds for all $x_0 \in \mathbb{R}^n$ where f is continuous.

General case: In the general case we cannot argue with continuity – but we will argue with convolution.

Let η be a typical bump function, $\eta \in C_c^\infty(B(0, 1))$, $\eta \geq 0$, $\int \eta = 1$ and set $\eta_\varepsilon := \varepsilon^{-n} \eta(\cdot/\varepsilon)$. Fix $x \in \mathbb{R}^n$. Then the assumption imply

$$f_\varepsilon(x) = \langle f, \eta_\varepsilon(x - \cdot) \rangle = 0.$$

That is $f_\varepsilon(x) = 0$ for all $x \in \mathbb{R}^n$. On the other hand we have by Corollary 4.33,

$$\|f\|_{L^1(K)} = \|0 - f\|_{L^1(K)} = \|f_\varepsilon - f\|_{L^1(K)} \xrightarrow{\varepsilon \to 0} = 0.$$

That is $f = 0$ $\mathcal{L}^n$-a.e. in K, and since K was arbitary $f = 0$ in a.e. in $\mathbb{R}^n$. □

Alright, so we can describe L^1_{loc}-functions as distributions. And all distributions have distributional derivatives.

What is the relation to Sobolev space? Here it is.

Theorem 4.46. *Let $p \in [1, \infty)$ and $f \in L^p(\mathbb{R}^n)$. The following are equivalent.*

(1) $f \in W^{1,p}(\mathbb{R}^n)$

(2) for each $\alpha \in \{1, \ldots, n\}$ *there exists* $g_\alpha \in L^p(\mathbb{R}^n)$ *such that the distributional derivative* $\partial_\alpha f$ *coincides with* g_α *in the following sense:*

$$\partial^\alpha f[\varphi] \equiv -\langle f, \partial^\alpha \varphi\rangle = \int_R \mathbb{R}^n g_\alpha \varphi.$$

We will simply write this as $\partial_\alpha f = g_\alpha$.

Proof. $\underline{\Rightarrow}$ Let $f \in W^{1,p}(\mathbb{R}^n)$. By definition there exists $f_k \in C_c^\infty(\mathbb{R}^n)$ which is a $W^{1,p}$-Cauchy sequence converging in $L^p(\mathbb{R}^n)$ to f.

Let $g_{\alpha,k} := \partial_\alpha f_k$. These are L^p-Cauchy sequences! So they have a limit $g_\alpha \in L^p(\mathbb{R}^n)$.

Now let $\varphi \in C_c^\infty(\mathbb{R}^n)$. We have from the integration by parts formula (for Riemann Integrals!)

$$\int_{\mathbb{R}^n} \partial_\alpha f_k \varphi = -\int_{\mathbb{R}^n} f_k \partial_\alpha \varphi$$

Taking the limit $k \to \infty$ we find

$$\int_{\mathbb{R}^n} g_\alpha \varphi = -\int_{\mathbb{R}^n} f \partial_\alpha \varphi = \partial_\alpha f[\varphi].$$

That proves the first direction.

$\underline{\Leftarrow}$ Now assume that

$$\int_{\mathbb{R}^n} g_\alpha \varphi = -\int_{\mathbb{R}^n} f \partial_\alpha \varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

Since g_α and f are L^p-maps we can approximate them with convolutions

$$g_{\alpha,\varepsilon} := g_\alpha * \eta_\varepsilon$$

and

$$f_\varepsilon := f * \eta_\varepsilon.$$

We may assume that $\eta(x) = \eta(-x)$, and then we have by Exercise 4.21

$$\int_{\mathbb{R}^n} (g_\alpha * \eta_\varepsilon)\varphi = \int_{\mathbb{R}^n} g_\alpha\,(\varphi * \eta_\varepsilon)$$

Since $(\varphi * \eta_\varepsilon) \in C_c^\infty(\mathbb{R}^n)$ we have

$$= -\int_{\mathbb{R}^n} f\,\partial_\alpha\,(\varphi * \eta_\varepsilon) = \int_{\mathbb{R}^n} \partial_\alpha(f * \eta_\varepsilon)\,\varphi.$$

That is, we have

$$\langle g_\alpha * \eta_\varepsilon, \varphi\rangle = \langle \partial_\alpha(f * \eta_\varepsilon), \varphi\rangle.$$

This holds for all $\varphi \in C_c^\infty(\mathbb{R}^n)$, so by the fundamental theorem of Calculus of Variations, Lemma 4.45,

$$g_\alpha * \eta_\varepsilon = \partial_\alpha(f * \eta_\varepsilon).$$

Thus we have

$$\|\partial_\alpha(f * \eta_\varepsilon) - g_\alpha\|_{L^p(\mathbb{R}^n)} = \|g_\alpha * \eta_\varepsilon - g_\alpha\|_{L^p(\mathbb{R}^n)} \xrightarrow{\varepsilon\to 0} 0.$$

That is $f * \eta_\varepsilon$ is a $W^{1,p}$-Cauchy sequence that converges to f.

However $f * \eta_\varepsilon \in C^\infty(\mathbb{R}^n)$, not in $C_c^\infty(\mathbb{R}^n)$. But this can be remedied easily.

Let $\chi \in C_c^\infty(B(0,2),[0,1])$, $\chi \equiv 1$ in $B(0,1)$ (this can be done similar to the bump functions). Set

$$f_k(x) := \chi(x/k)\ (f * \eta_{1/k})(x).$$

We have

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \leq \|(1-\chi(x/k))f\|_{L^p(\mathbb{R}^n)} + \|\chi(x/k)\left(f - (f * \eta_{1/k})\right)\|_{L^p(\mathbb{R}^n)}.$$

The first term tends to zero by dominated convergence, the second by the L^p-convergence of the convolution.

$$\begin{aligned}\|\partial_\alpha f_k - g_\alpha\|_{L^p(\mathbb{R}^n)} \leq &\|(1-\chi(x/k))g_\alpha\|_{L^p(\mathbb{R}^n)} + \|\chi(x/k)\left(g_\alpha - \partial_\alpha(f * \eta_{1/k})\right)\|_{L^p(\mathbb{R}^n)}\\ &+ \|\partial_\alpha(\chi(x/k))\|_{L^\infty}\|(f * \eta_{1/k})(x) - f\|_{L^p(\mathbb{R}^n)}\end{aligned}$$

Observe that $|\partial_\alpha \chi(x/k)| \leq C/k \xrightarrow{k\to\infty} 0$. So everything converges.

□

So for now we take with us: Sobolev functions $f \in W^{1,p}(\mathbb{R}^n)$ are maps $f \in L^p(\mathbb{R}^n)$ such that $\partial_\alpha f \in L^p(\mathbb{R}^n)$ (where ∂_α denotes the distributional derivative).

Let us stress that distributional derivatives are not the same as a.e. derivatives. "a.e. derivatives are not a good object, since they "forget" things.

Example 4.47. Let h be the *Heaviside function*

$$H(x) := \begin{cases} 1 & x > 0\\ 0 & x \leq 0.\end{cases}$$

Then clearly for almost all $x \in \mathbb{R}$ we have $H'(x)$ exists and $H'(x) = 0$ a.e. – however H is not a constant map.

What is the distributional derivative? By definition for $\varphi \in C_c^\infty(\mathbb{R})$,

$$H'[\varphi] := -\int_{\mathbb{R}} H(x)\varphi'(x)dx = -\int_0^\infty \varphi'(x)dx = \varphi(0) - \varphi(\infty) = \varphi(0).$$

(Here we can use the Riemann integral since φ is continuous).

What does that tell us? Well if we denote δ_0 again the Dirac delta, then

$$\varphi(0) = \int_{\mathbb{R}} \varphi(x)d\delta_0(x).$$

That is we have

$$H' = \delta_0$$

in distributional sense.

Proposition 4.48. *Let $1 \leq p < \infty$ and assume that $f \in W^{1,p}(\mathbb{R}^n)$ satisfies $\partial_\alpha f = 0$ for all $\alpha = 1, \ldots, n$. Then f is constant a.e., that is for some $c \in \mathbb{R}$ we have*

$$f(x) = c \quad \text{a.e. } x \in \mathbb{R}^n$$

(actually $c = 0$ since any constant function in $L^p(\mathbb{R}^n)$ is zero).

Proof. By definition there must be $f_k \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|f_k - f\|_{L^p(\mathbb{R}^n)} + \|Df_k\|_{L^p(\mathbb{R}^n)} < \frac{1}{k}$$

By the classical fundamental theorem we have

$$f_k(x) - f_k(y) = \int_0^1 \sum_{\alpha=1}^n \partial_\alpha f_k(x + t(y-x))(y-x)^\alpha \, dt.$$

Integrating this inequality in both x and y in $B(0,R)$ (for some fixed R) we have

$$\begin{aligned}
&\int_{B(0,R)} \int_{B(0,R)} |f_k(x) - f_k(y)| dx dy \\
\leq &2R \left(\int_0^{\frac{1}{2}} \int_{B(0,R)} \int_{B(0,R)} |Df_k(x + t(y-x))| dx dy dt + \int_{\frac{1}{2}}^1 \int_{B(0,R)} \int_{B(0,R)} |Df_k(x + t(y-x))| dy \, dx dt \right)
\end{aligned}$$

Now for $t \in (0, \frac{1}{2})$ by the substitution rule (we use also the convexity of $B(0,R)$) and Hölder's inequality

$$\int_{B(0,R)} |Df_k(x+t(y-x))| dx \leq \frac{1}{(1-t)^n} \int_{B(0,R)} |Df_k(z))| dz \leq 2^n \left(\int_{B(0,R)} |Df_k(z))|^p dz \right)^{\frac{1}{p}} (\mathcal{L}^n(B(0,R)))^{1-\frac{1}{p}}$$

Doing a similar argument for the $\int_{\frac{1}{2}}^1$-integral we obtain

$$\int_{B(0,R)} \int_{B(0,R)} |f_k(x) - f_k(y)| dx dy \leq C(R) \|Df_k\|_{L^p(B(0,R))}.$$

Now since f_k converges to f in $L^p(\mathbb{R}^n)$ we have that f_k converges to f in $L^1(B(0,R))$ (Hölder inequality again!), so by taking the limit we find

$$\int_{B(0,R)} \int_{B(0,R)} |f(x) - f(y)| dx dy \leq \lim_{k \to \infty} C(R) \|Df_k\|_{L^p(B(0,R))} = 0.$$

Consequently, $|f(x) - f(y)| = 0$ for $\mathcal{L}^{2n}$-a.e. (x,y). By Fubini's theorem (e.g. the version in Proposition 4.6)there must be some $x \in B(0,R)$ such that $f(y) = f(x)$ for $\mathcal{L}^n$-a.e. $y \in B(0,R)$. Setting $c := f(x)$ we find that $f(y) = c$ a.e., that is f is constant in $B(0,R)$. This holds for any $R > 0$ so $f = c$ a.e. in $\mathbb{R}^n$. □

Exercise 4.49. *Use Proposition 4.48 to show that $\chi_{[0,1]} \notin W^{1,p}(\mathbb{R})$ for any $p \in [1, \infty)$, where as usual*

$$\chi_{[0,1]}(x) = \begin{cases} 1 & x \in [0,1], \\ 0 & x \in \mathbb{R} \setminus [0,1]. \end{cases}$$

More precisely show that

(1) $\chi_{[0,1]} \in L^p(\mathbb{R})$ *for any* $p \in [1,\infty]$
(2) If there was $f \in L^1_{loc}(\mathbb{R})$ *such that*

$$\chi_{[0,1]}[-\varphi'] = -\int f(x)\varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R})$$

then $f(x) = 0$ *for* $\mathcal{L}^1$*-a.e.* $x \in \mathbb{R}$ *(Hint: Fundamental theorem: Lemma 4.45)*
(3) Conclude with Proposition 4.48.

We will treat Sobolev functions in way more detail later (next semester), let us just give one more application of the approximation.

Lemma 4.50. *Let* $f \in W^{1,1}(\mathbb{R})$ *(1 dimension only!) then* f *is continuous. That is, there exists* $\bar{f} : \mathbb{R} \to \mathbb{R}$ *continuous such that* $f = \bar{f}$ *a.e.*

Proof. Let $f_k \in C_c^\infty(\mathbb{R})$ be the approximation of f. By the fundamental theorem we have

$$f_k(x) - f_k(y) = \int_x^y f_k'(z)dz.$$

That is

$$|f_k(x) - f_k(y)| \leq \int_{[x,y]} |f_k'(z)|dz.$$

Restricting ourselves to a ball $[-R,R]$ we can apply Vitali's convergence theorem, Theorem 3.59: since f_k' converges in $L^1[-R,R]$ to f' we have that the integral is uniformly absolutely continuous. That is, for any $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x, y \in [-R,R]$ with $|x-y| < \delta$ we have

$$\int_{[x,y]} |f_k'(z)|dz < \varepsilon.$$

Thus for any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f_k(x) - f_k(y)| < \varepsilon \quad x, y \in [-R,R] : \quad |x-y| < \delta.$$

This is uniformly equicontinuity of $(f_k)_{k=1}^\infty : [-R,R] \to \mathbb{R}$. If only we had uniform boundedness, then we could use Arzela-Ascoli!

But don't despair. Since f_k converges to f in $L^1([-R,R])$, by Theorem 3.51 there exists a subsequence f_{k_i} which converges a.e. to f in $[-R,R]$. Let $x_0 \in [-R,R]$ be a point such that $f_{k_i}(x_0) \to f(x_0)$. Then we have from the equicontinuity

$$|f_{k_i}(x) - f_{k_i}(x_0)| \leq C,$$

or rather

$$|f_{k_i}(x)| \leq C + \sup_i |f_{k_i}(x_0)| < \infty \quad \forall x \in [-R,R].$$

This is uniform boundedness, so by Arzela-Ascoli, taking yet another subsequence if necessary we conclude that $f_{k_{i_j}}$ uniformly converges to a continuous limit function $g : [-R,R] \to \mathbb{R}$. Since on the other hand $f_{k_{i_j}}$ converges a.e. to f we have that $f = g$ a.e. in $B(0,R)$.

Since the *continuous representative*[20] of a Lebesgue-integrable function is unique, Exercise 4.51, we can let $R \to \infty$ to find a continuous representative of f. $\square$

Exercise 4.51. *Let $f \in L^1_{loc}(\Omega)$ for some open set $\Omega \subset \mathbb{R}^n$ (and the Lebesgue measure). Assume there exist $g, h : \Omega \to \mathbb{R}$ which are continuous and $f = g$ a.e. and $f = h$ a.e. Show that $g = h$ everywhere.*

Exercise 4.52. *Use Lemma 4.50 to essentially reprove Exercise 4.52: show that $\chi_{[0,1]} \notin W^{1,p}(\mathbb{R})$ for any $p \in [1, \infty)$, where as usual*

$$\chi_{[0,1]}(x) = \begin{cases} 1 & x \in [0,1], \\ 0 & x \in \mathbb{R} \setminus [0,1]. \end{cases}$$

5. Differentiation of Radon measures - Radon-Nikodym Theorem on $\mathbb{R}^n$

5.1. **Preparations: Besicovitch Covering theorem.** If a set A is covered by closed balls, we would sometimes like it to be covered by countably many *disjoint* closed balls (e.g. so we can sum of $\mu(B)$). One very useful theorem is the

Theorem 5.1 (Besicovitch Covering theorem)**.** *There exists a constant N_n depending only on the dimension n with the following property.*

If $\mathcal{F}$ is a family of closed balls $B(x, r)$, $r > 0$ in $\mathbb{R}^n$ with finite maximal radius, i.e.

$$\sup \{\operatorname{diam} B : \quad B \in \mathcal{F}\} < \infty,$$

and if A is the set of centers of balls in $\mathcal{F}$, then there exist N_n countable families $\mathcal{G}_1, \dots, \mathcal{G}_{N_n}$ of balls where each family $\mathcal{G}_i \subset \mathcal{F}$ consists of pairwise disjoint balls in $\mathcal{F}$ such that

$$A \subset \bigcup_{i=1}^{N_n} \bigcup_{B \in \mathcal{G}_i} B.$$

Proof. The proof is lengthy and technical and we skip it here. See [Evans and Gariepy, 2015, Theorem 1.27]. $\square$

As an application of Besicovitch covering theorem we have

Theorem 5.2. *Let μ be a Borel measure on $\mathbb{R}^n$ and $\mathcal{F}$ any collection of closed balls $\overline{B(a,r)}$ for $a \in \mathbb{R}^n$ and $r \in (0, \infty)$.*

Let A denote the set of centers of the balls in $\mathcal{F}$ and assume $\mu(A) < \infty$.

Moreover assume that for each $a \in A$

$$\inf\{r : \quad \overline{B(a,r)} \in \mathcal{F}\} = 0.$$

[20] more on this later: see Definition 5.17

Then for each open set $U \subset \mathbb{R}^n$ there exists a countable collection $\mathcal{G}$ of disjoint balls in $\mathcal{F}$ such that

$$\bigcup_{B\in\mathcal{G}} B \subset U$$

and

$$\mu\left((A\cap U)\setminus \bigcup_{B\in\mathcal{G}} B\right) = 0.$$

(observe there are no assumptions on the measurability of A).

Proof. See [Evans and Gariepy, 2015, Theorem 1.28]. $\square$

5.2. **The Radon-Nikodym Theorem.** In this section[21] let ν and μ be *always* two Radon measures on $\mathbb{R}^n$. *Extremely formally* we could hope to write

$$\mu(A) = \int_A d\mu = \int_A \frac{d\mu}{d\nu} d\nu. \tag{5.1}$$

And then (as we do in calculus), we could interpret $\frac{d\mu}{d\nu}$ as the derivative of μ in direction ν

$$\frac{d\mu}{d\nu} = D_\mu \nu.$$

In some sense (5.1) is the fundamental theorem of calculus for measures.

Definition 5.3. For each point $x \in \mathbb{R}^n$ set

$$\overline{D}_\mu\nu(x) := \begin{cases} \limsup_{r\to 0} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} & \text{if } \mu(\overline{B(x,r)}) > 0 \text{ for all } r>0\\ +\infty & \text{if } \mu(\overline{B(x,r)}) = 0 \text{ for some } r>0\end{cases}$$

and

$$\underline{D}_\mu\nu(x) := \begin{cases} \liminf_{r\to 0} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} & \text{if } \mu(\overline{B(x,r)}) > 0 \text{ for all } r>0\\ +\infty & \text{if } \mu(\overline{B(x,r)}) = 0 \text{ for some } r>0\end{cases}$$

If $\overline{D}_\mu\nu(x) = \underline{D}_\mu\nu(x) < \infty$ we say that ν is *differentiable with respect to μ*. We also call $D_\mu\nu$ the *density* of ν with respect to μ. [22]

Example 5.4.
- If $\nu(B) := \lambda\mu(B)$ then $D_\mu\nu(x) = \lambda$.

---

[21]We follow to a substantial extend [Evans and Gariepy, 2015].

[22]This indeed has some features of a derivative, because we could believe that for "a.e." x, $\nu(\overline{B(x,0)}) = \mu(\overline{B(x,0)}) = 0$ (where $B(x,0) = \{x\}$) and then write

$$\liminf_{r\to 0} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} = \liminf_{r\to 0} \frac{\nu(\overline{B(x,r)}) - \nu(\overline{B(x,0)})}{\mu(\overline{B(x,r)}) - \mu(\overline{B(x,0)})}$$

- We will later see (Theorem 5.15, as a consequence of the Radon-Nikodym theorem Theorem 5.13) that if for a μ-integrable function f
$$\nu = f \llcorner \mu$$
i.e. for μ-measurable A,
$$\nu(A) := \int_A f d\mu$$
then
$$D_\mu \nu(x) = f(x) \quad \mu - \text{a.e.}.$$
This is easy to see for $\mu = \mathcal{L}^n$ and f continuous in x:
Then
$$\nu(\overline{B(x,r)}) = \int_{\overline{B(x,r)}} f d\mu = \mu(\overline{B(x,r)}) f(x) + \int_{\overline{B(x,r)}} (f - f(x))\, d\mu.$$
By continuity for any $\varepsilon > 0$ there exists $r > 0$ such that $\sup_{z \in B(x,r)} |f(z) - f(x)| < \varepsilon$, so we have
$$\left|\nu(\overline{B(x,r)}) - \mu(\overline{B(x,r)}) f(x)\right| \leq \varepsilon \mu(\overline{B(x,r)})$$
So we have
$$\left|\frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} - f(x)\right| \xrightarrow{r \to 0} 0.$$

Our first goal is to understand when $D_\mu \nu$ exists and when (5.1) holds.

For this the following observation is useful

Lemma 5.5. *Let μ be a Radon measure on $\mathbb{R}^n$ and ν be any other measure on $\mathbb{R}^n$ (not necessarily Borel regular or anything).*

Assume that $A \subset \mathbb{R}^n$ is a (possibly not measurable) set such that
$$\nu(A \cap V) \leq \mu(A \cap V) \quad \textit{for all bounded Borel sets } V \subset \mathbb{R}^n$$
Then
$$\nu(A) \leq \mu(A).$$

Proof. If $\mu(A) = \infty$ there is nothing to show. So assume $\mu(A) < \infty$. Then there must be an open set $U \supset A$ such that
$$\mu(U) \leq \mu(A) + \varepsilon.$$
We can write $\mathbb{R}^n = \bigcup_{i=1}^\infty V_i$ where each V_i is a bounded Borel set and $V_i \cap V_j = \emptyset$ and thus $\mu(V_i) < \infty$ for $i \neq j$ (e.g. take partially open and closed cubes).

We then have
$$\nu(A) \leq \sum_{i=1}^\infty \nu(A \cap V_i) \leq \sum_{i=1}^\infty \mu(A \cap V_i) \leq \sum_{i=1}^\infty \mu(U \cap V_i) = \mu(U).$$

In the last step we used the measurability of $U \cap V_i$ (but A may not be measurable!) Thus,

$$\nu(A) \leq \mu(A) + \varepsilon.$$

This holds for any $\varepsilon > 0$ and we can conclude by letting $\varepsilon \to 0$. $\square$

Lemma 5.6. *Let μ, ν be Radon measures on $\mathbb{R}^n$. Fix $0 < \alpha < \infty$. Let $A \subset \mathbb{R}^n$*

(1) If $A \subset \{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) \leq \alpha\}$ then $\nu(A) \leq \alpha\mu(A)$
(2) If $A \subset \{x \in \mathbb{R}^n : \quad \overline{D}_\mu \nu(x) \geq \alpha\}$ then $\nu(A) \geq \alpha\mu(A)$

Observe that we don't assume A to be ν or μ measurable!

Proof. (1) Fix $\varepsilon > 0$ and assume $A \subset \{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) \leq \alpha\}$.
Let $V \subset \mathbb{R}^N$ be any bounded Borel set (in particular $\mu(V), \nu(V) < \infty$).
We are going to show

$$\nu(A \cap V) \leq (\alpha + \varepsilon)\mu(A \cap V). \tag{5.2}$$

Since this holds for any bounded Borel set V, in view of Lemma 5.5, we find $((\alpha+\varepsilon)\mu$ is still a Radon measure!)

$$\nu(A) \leq (\alpha + \varepsilon)\mu(A).$$

Letting $\varepsilon \to 0$ we can then conclude.

Take any bounded open set $U \subset \mathbb{R}^n$ with $A \cap V \subset U$ (this exists, because V is bounded). Then $\nu(U) < \infty$.

We define a family of balls:

$$\mathcal{F} := \left\{\overline{B(a,r)} : \quad \text{where } a \in A \cap V, r > 0 : \ \overline{B(a,r)} \subset U, \nu(\overline{B(a,r)}) \leq (\alpha + \varepsilon)\mu(\overline{B(a,r)})\right\}$$

Since for each $a \in A$ we have

$$\liminf_{r \to 0} \frac{\nu(\overline{B(a,r)})}{\mu(\overline{B(a,r)})} \leq \alpha$$

we find that in particular for each $a \in A \cap V$

$$\inf\{r > 0 : \quad \overline{B(a,r)} \in \mathcal{F}\} = 0.$$

By Theorem 5.2 there is a countable subfamily $\mathcal{G} \subset \mathcal{F}$ of pairwise disjoint balls such that

$$\bigcup_{B \in \mathcal{G}} B \subset U$$

and

$$\nu((A \cap V) \cap U \setminus \bigcup_{B \in \mathcal{G}} B) = \nu((A \cap V) \setminus \bigcup_{B \in \mathcal{G}} B) = 0$$

Then, for each $B \in \mathcal{G} \subset \mathcal{F}$

$$\nu(B) \leq (\alpha + \varepsilon)\mu(B)$$

and summing this up over $B \in \mathcal{G}$ (observe the B are all measurable as closed balls) we have

$$\nu(A\cap V) = \nu(A\cap V\cap U) \leq \sum_{B\in\mathcal{G}} \nu(B)+0 \leq \sum_{B\in\mathcal{G}} (\alpha+\varepsilon)\mu(B)+0 = (\alpha+\varepsilon)\mu(\bigcup_{B\in\mathcal{G}} B) \leq (\alpha+\varepsilon)\mu(U).$$

We have this inequality for any bounded open set $U \subset \mathbb{R}^n$ with $A \cap V \subset U$. Taking the infimum over all such U we find in view of Theorem 1.68 Equation (5.2) holds and we can conclude.

(2) (similar)

□

Theorem 5.7 (Differentiating measures)**.** *Let μ and ν be Radon measures on $\mathbb{R}^n$. Then*

(1) $D_\mu\nu(x)$ exists and is finite for μ-a.e.[23] *$x \in \mathbb{R}^n$ and*
(2) $D_\mu\nu$ is μ-measurable.

We may assume that $\mu(\mathbb{R}^n)$ and $\nu(\mathbb{R}^n) < \infty$. Otherwise we consider $\mu\llcorner K$ and $\nu\llcorner K$ for compact sets and show (1) and (2) in open sets $U \subset K$ with $\operatorname{dist}(U, K) > 0$. Then we can argue by exhaustion.

We split the proof of Theorem 5.7 into different parts.

Lemma 5.8. *Assumptions as in Theorem 5.7. Then $D_\mu\nu$ exists and is finite μ-a.e.*

Proof. Let $A := \{x \in \mathbb{R}^n : \overline{D}_\mu\nu(x) = \infty\}$. Then for each $\alpha > 0$

$$A \subset \{x \in \mathbb{R}^n : \overline{D}_\mu\nu(x) \geq \alpha\},$$

and we can apply Lemma 5.6(2) to obtain

$$\mu(A) \leq \frac{1}{\alpha}\nu(A).$$

If $\nu(A) < \infty$ we can let $\alpha \to \infty$ to conclude $\mu(A) = 0$. If $\nu(A) = \infty$ then since $\nu(A \cap (B(0,R))) < \infty$ (ν is Radon) we find $\mu(A \cap B(0,R)) = 0$ for all $R > 0$ and thus again $\mu(A) = 0$.

That is

$$\overline{D}_\mu\nu(x) < \infty \quad \mu\text{-a.e.}$$

Now let $0 < a < b$ and set

$$R(a,b) := \{x \in \mathbb{R}^n : \underline{D}_\mu\nu(x) < a < b < \overline{D}_\mu\nu(x) < \infty\}.$$

We apply again Lemma 5.6 and have

$$b\mu(R(a,b)) \leq \nu(R(a,b)) \leq a\mu(R(a,b)).$$

[23] this μ-a.e. takes care of pathological examples. E.g. assume $\mu(\mathbb{R}^n) = 0$ then $D_\mu\nu \equiv \infty$, but still $D_\mu\nu(x) = 0$ for μ-a.e. $x \in \mathbb{R}^n$ since $\mu(\mathbb{R}^n) = 0$

Since $b > a$ and (w.l.o.g. otherwise again intersect with balls) $\mu(R(a,b)) < \infty$ we conclude $\mu(R(a,b)) = 0$. Thus

$$\left\{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) < \overline{D}_\mu \nu(x) < \infty\right\} = \bigcup_{a \in \mathbb{Q}} \bigcup_{b \in \mathbb{Q} \cap (a,\infty)} R(a,b).$$

The right-hand side is a countable union of μ-zerosets, so

$$\mu\left(\left\{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) < \overline{D}_\mu \nu(x) < \infty\right\}\right) = 0.$$

Since we always have $\underline{D}_\mu \nu(x) \geq \overline{D}_\mu \nu(x)$ we conclude that

$$\mu\left(\left\{x \in \mathbb{R}^n : \quad \underline{D}_\mu \nu(x) \neq \overline{D}_\mu \nu(x)\right\}\right) = 0.$$

That is $\underline{D}_\mu \nu(x) \neq \overline{D}_\mu \nu(x) < \infty$ for μ-a.e. x. □

One ingredient is the upper semicontinuity of $x \mapsto \mu(B(x,r))$.

Lemma 5.9. *Let μ be a Radon measure, $x \in \mathbb{R}^n$ and $r > 0$. Then*

$$\limsup_{y \to x} \mu(\overline{B(y,r)}) \leq \mu(\overline{B(x,r)})$$

Proof. Since $\overline{B(x,r)}$ is Borel, it is measurable and $\chi_{\overline{B(x,r)}}$ is μ-integrable since $\overline{B(x,r)}$ is compact.

Let $y_k \to y$ and set $f_k := \chi_{\overline{B(y_k,r)}}$ and $f := \chi_{\overline{B(x,r)}}$. While it is not so clear whether $f_k \to f$ everywhere we certainly have for any $z \in \mathbb{R}^n$

$$\limsup_{k \to \infty} f_k(z) \leq f(z) \tag{5.3}$$

Indeed, the only case we need to consider is $z \in \mathbb{R}^n$ with $f(z) = 0$, otherwise the inequality is obvious. But then $z \in \mathbb{R}^n \backslash \overline{B(x,r)}$, and thus $\operatorname{dist}(z, \overline{B(x,r)}) > \delta$ for some $\delta > 0$. Now if $|y_k - x| < \frac{\delta}{2}$ then $z \not\in B(y_k, r)$, so since $y_k \to x$ we have $f_k(z) = 0$ for all but finitely many $k \in \mathbb{N}$. So (5.3) is established.

From (5.3) we conclude

$$\liminf_{k \to \infty} (1 - f_k) \geq (1 - f).$$

Applying Fatou's lemma Corollary 3.9 on $X = B(x, 2r)$

$$\int_{B(x,2r)} (1-f) d\mu \leq \int_{B(x,2r)} \liminf_{k \to \infty} (1 - f_k) d\mu \leq \liminf_{k \to \infty} \int_{B(x,2r)} (1 - f_k) d\mu$$

That is

$$\underbrace{\mu(B(x,2r))}_{<\infty} - \mu(B(x,r)) \leq \liminf_{k \to \infty} \left(\mu(B(x,2r)) - \mu(B(y_k,r))\right).$$

Subtracting the constant $\mu(B(x,2r))$ from both sides we conclude. □

Lemma 5.10. *Assumptions as in Theorem 5.7. Then $D_\mu \nu$ is μ-measurable.*

Proof. By Lemma 5.9, the functions $x \mapsto \mu(\overline{B(x,r)})$ and $x \mapsto \nu(\overline{B(x,r)})$ are upper semi-continuous. In view of Exercise 2.16 both functions are then μ-measurable. Consequently for each $r > 0$

$$f_r(x) := \begin{cases} \frac{\nu(\overline{B(x,r)})}{\mu(\overline{B(x,r)})} & \text{if } \mu(\overline{B(x,r)}) > 0 \\ +\infty & \text{if } \mu(\overline{B(x,r)}) = 0 \end{cases}$$

is μ-measurable. Since by Lemma 5.8 $D_\mu \nu(x) = \lim_{r\to 0} f_r(x)$ for μ-a.e. we have by Theorem 2.13 that $D_\mu \nu$ is μ-measurable. □

Exercise 5.11. *Let μ, ν_1, ν_2 be Radon measures. Show that*

$$D_\mu(\nu_1 + \nu_2) = D_\mu \nu_1 + D_\mu \nu_2 \quad \mu\text{-a.e.}$$

We recall the definition of absolute continuity from Definition 3.54 (see also Lemma 3.55) and somewhat (as we shall see) define the opposite: mutually singular.

Definition 5.12. Assume μ and ν are Borel measures on $\mathbb{R}^n$.

(1) The measure ν is *absolutely continuous* with respect to μ, in symbols $\nu \ll \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$ for all $A \subset \mathbb{R}^n$.
(2) The measures ν and μ are *mutually singular*, in symbols $\nu \perp \mu$ if there exists a Borel set $B \subset \mathbb{R}^n$ such that

$$\mu(\mathbb{R}^n \setminus B) = \nu(B) = 0.$$

Here is the fundamental theorem for measures, also called the *Radon-Nikodym Theorem*.

Theorem 5.13. *Let ν, μ be Radon measures on $\mathbb{R}^n$ with $\nu \ll \mu$.*

Then for all μ-measurable sets A we have

$$\nu(A) = \int_A D_\mu \nu \, d\mu.$$

Proof. Let $A \subset \mathbb{R}^n$ be μ-measurable. Then, by Lemma 3.55, A is also ν-measurable.

Set

$$Z := \{x \in \mathbb{R}^n : \quad D_\mu \nu(x) = 0\}$$

and

$$I := \{x \in \mathbb{R}^n : \quad D_\mu \nu(x) = +\infty\}$$

By Theorem 5.7 $\mu(I) = 0$ and thus since $\nu \ll \mu$ we have $\nu(I) = 0$. Also $\nu(I) = 0$ by Lemma 5.6. In particular

$$\nu(Z) = 0 = \int_Z D_\mu \nu \, d\mu$$

$$\nu(I) = 0 = \int_I D_\mu \nu \, d\mu.$$

Now let A be μ-measurable and fix $t \in (1, \infty)$ and set for $m \in \mathbb{Z}$

$$A_m := A \cap \{x \in \mathbb{R}^n : \quad t^m \leq D_\mu \nu(x) < t^{m+1}\}.$$

Then A_m is also μ-measurable, and thus ν-measurable, Lemma 3.55, and we have

$$t^m \mu(A_m) \leq \int_{\{x:\, t^m \leq D_\mu \nu(x)\}} t^m d\mu \leq \int_{A_m} D_\mu \nu d\mu$$

Since

$$A \setminus \bigcup_{m \in \mathbb{Z}} A_m \subset Z \cup I \cup \{x \in \mathbb{R}^n : \quad \overline{D}_\mu \nu(x) \neq \underline{D}_\mu \nu(x)\}$$

we have that $\mu(A \setminus \bigcup_{m \in \mathbb{Z}} A_m) = 0$ and thus ($(A_m)_{m \in \mathbb{Z}}$ are pairwise disjoint!)

$$\sum_{m \in \mathbb{Z}} t^m \mu(A_m) \leq \sum_m \int_{A_m} D_\mu \nu d\mu = \int_A D_\mu \nu \; d\mu.$$

Since $\nu \ll \mu$ we also have $\nu(A \setminus \bigcup_{m \in \mathbb{Z}} A_m) = 0$ and thus

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m).$$

With the help of Lemma 5.6 we find

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m) \leq \sum_m t^{m+1} \mu(A_m) \leq t \int_A D_\mu \nu \; d\mu$$

and similarly

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m) \geq \sum_m t^m \mu(A_m) = t^{-1} \sum_m t^{m+1} \mu(A_m) \geq t^{-1} \int_A D_\mu \nu \; d\mu.$$

That is for any $t > 1$ we have

$$t^{-1} \int_A D_\mu \nu \; d\mu \leq \nu(A) \leq t \int_A D_\mu \nu \; d\mu.$$

Letting $t \to 1$ we conclude. □

If $\nu \not\ll \mu$ we can get also the following refinement

Theorem 5.14 (Lebesgue Decomposition Theorem)**.** *Let ν and μ be Radon measures on $\mathbb{R}^n$.*

(1) Then we can decompose

$$\nu = \nu_{ac} + \nu_s$$

where the absolutely continuous part $\nu_{ac} \ll \mu$ and the singular part $\nu_s \perp \mu$.

(2) Furthermore,

$$D_\mu \nu = D_\mu \nu_{ac}, \quad D_\mu \nu_s = 0 \quad \mu\text{-a.e.}$$

and consequently

$$\nu(A) = \int_A D_\mu \nu d\mu + \nu_s(A)$$

for each Borel set $A \subset \mathbb{R}^n$[24].

Proof. Again assume $\mu(\mathbb{R}^n), \nu(\mathbb{R}^n) < \infty$ or argue with compact exhaustion.

We begin by constructing the singular measure. For this we are going to find a suitable Borel set B and set

$$\nu_{ac} := \nu \llcorner B, \quad \nu_s := \nu \llcorner (\mathbb{R}^n \setminus B).$$

To find B define $\mathcal{E}$ the set of "good candidates", i.e. Borel sets A where $\mu(\mathbb{R}^n \setminus A) = 0$, i.e.

$$\mathcal{E} := \{A \subset \mathbb{R}^n : \quad A \text{ Borel}, \, \mu(\mathbb{R}^n \setminus A) = 0\}.$$

We are going to choose B such that

$$\nu(B) = \min_{A \in \mathcal{E}} \nu(A).$$

But for this we have to show that such a B exists. Certainly, $\mathcal{E} \neq \emptyset$ so there must be some $B_k \in \mathcal{E}$ with

$$\nu(B_k) \leq \inf_{A \in \mathcal{E}} \nu(A) + \frac{1}{k} \quad k = 1, \ldots$$

Write $B := \bigcap_{k=1}^{\infty} B_k$. Then

$$\mu(\mathbb{R}^n \setminus B) \leq \sum_{k=1}^{\infty} \mu(\mathbb{R}^n \setminus B_k) = 0.$$

That is $B \in \mathcal{E}$ and thus

$$\nu(B) = \inf_{A \in \mathcal{E}} \nu(A). \tag{5.4}$$

Define

$$\nu_{ac} := \nu \llcorner B$$

and

$$\nu_s := \nu \llcorner (\mathbb{R}^n \setminus B).$$

Both, ν_{ac} and ν_s are Radon measures, see Exercise 1.67.

We now show that $\nu_{ac} \ll \mu$. Assume $A \subset \mathbb{R}^n$ with $\mu(A) = 0$. Since μ is a Borel regular measure we may assume that A is Borel (otherwise we pass to $A \subset \tilde{A}$ with $\mu(\tilde{A}) = \mu(A)$). Then $\nu_{ac}(A) = \nu(A \cap B)$. Observe that

$$\mu(\mathbb{R}^n \setminus (B \setminus A)) \leq \mu(\mathbb{R}^n \setminus B) + \mu(A) = 0,$$

so $B \setminus A \in \mathcal{E}$. Then by (5.4) (observe both A and B are Borel, so ν-measurable)

$$\nu(B) \leq \nu(B \setminus A) = \nu(B) - \nu(A \cap B).$$

Thus $\nu(A \cap B) = 0$, i.e. $\nu_{ac}(A) = 0$. I.e., we have established $\nu_{ac} \ll \mu$.

Now fix $\alpha > 0$ and set

$$C_\alpha := \{x \in B : \quad D_\mu \nu_s(x) \geq \alpha\}$$

[24]so A is μ and ν-measurable

In view of Lemma 5.6 we have

$$\alpha\mu(C_\alpha) \leq \nu_s(C_\alpha) = \nu(C_\alpha \cap (\mathbb{R}^n \setminus B)) = \nu(\emptyset) = 0.$$

That is $\mu(C_\alpha) = 0$ for all $\alpha > 0$, that is $D_\mu\nu_s(x) = 0$ for μ-a.e. $x \in B$ (and since $D_\mu\nu_s(x) = D_\mu(\nu \llcorner B)$ we have that $D_\mu\nu_s(x) = 0$ for μ-a.e. $x \in \mathbb{R}^n \setminus B$ by Theorem 5.7(1). So $D_\mu\nu_s(x) = 0$ for a.e. $x \in \mathbb{R}^n$, and thus (with the help of Exercise 5.11)

$$D_\mu\nu_{ac} = D_\mu\nu \quad \mu\text{-a.e.} \tag{5.5}$$

By the first Radon-Nikodym theorem, Theorem 5.13 we then have for any Borel set (thus ν_{ac}-measurable)

$$\nu_{ac}(A) = \int_A D_\mu\nu_{ac}d\mu \overset{(5.5)}{=} \int_A D_\mu\nu d\mu$$

Thus, by the definition of ν_s and ν_{ac}, for any Borel measure

$$\nu(A) = \nu_{ac}(A) + \nu_s(A) = \int_A D_\mu\nu d\mu + \nu_s(A)$$

□

5.3. **Lebesgue differentiation theorem.** The *integral average* of f over a measurable set E of finite, strictly positive measure will be denoted by

$$f_E \equiv ⨍_E d\mu := \mu(E)^{-1}\int_E f d\mu$$

Theorem 5.15. *Let μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^1_{loc}(\mathbb{R}^n, \mu)$. Then for μ-a.e. $x \in \mathbb{R}^n$*

$$\lim_{r\to 0} ⨍_{B(x,r)} f\, d\mu = f(x)$$

Exercise 5.16. *Show Theorem 5.15 for continuous f (using only continuity, no deep theorem).*

Proof of Theorem 5.15. For μ-measurable sets E set

$$\nu_+(E) := \int_E f_+ d\mu$$

and

$$\nu_-(E) := \int_E f_- d\mu.$$

Since $f \in L^1_{loc}(\mathbb{R}^n)$ we see that $\nu_\pm$ extend to Radon measures, and $\nu_\pm \ll \mu$.

By the Radon-Nikdoym theorem, Theorem 5.13

$$\nu_+(A) = \int_A D_\mu\nu_+ d\mu = \int_A f_+ d\mu \quad \text{for all } \mu\text{-measurable } A \subset \mathbb{R}^n$$

$$\nu_-(A) = \int_A D_\mu\nu_- d\mu = \int_A f_- d\mu \quad \text{for all } \mu\text{-measurable } A \subset \mathbb{R}^n$$

Since ν_+, ν_- are Radon measures, we see that $D_\mu \nu_\pm$ are $L^1_{loc}(d\mu)$-functions, so by Lemma 3.22(2), $f_+ = D_\mu \nu_+$ and $f_- = D_\mu \nu_-$ μ-a.e. in $\mathbb{R}^n$.

By the definition of the Radon-Nikodym derivative we have

$$\begin{aligned}\fint_{B(x,r)} f d\mu =& \frac{\nu_+(B(x,r))}{\mu(B(x,r))} - \frac{\nu_-(B(x,r))}{\mu(B(x,r))}\\ &\xrightarrow{r\to 0^+} D_\mu \nu^+(x) - D_\mu \nu^-(x) = f^+(x) - f^-(x) = f(x),\end{aligned}$$

which holds for μ-a.e. x in $\mathbb{R}^n$ – we can pass to $\overline{B(x,r)}$ by Proposition 1.72. □

Definition 5.17. Let $f \in L^1_{loc}(\mu)$.

The *precise representative* of f is the (pointwise!, not a.e., not a class) defined function

$$f^*(x) := \begin{cases} \lim_{r\to 0^+} \fint_{B(x,r)} f\, d\mu & \text{when the limit exists}\\ 0 & \text{otherwise}\end{cases}$$

By Theorem 5.15 the limit exists μ-a.e.

We say that $x \in \mathbb{R}^n$ is a *Lebesgue point* of f if

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) = 0.$$

Theorem 5.18. *Let μ be a Radon measure and $f \in L^1_{loc}(\mu)$. Then μ-a.e. point $x \in \mathbb{R}^n$ is a Lebesgue point of f*

Proof. By Theorem 5.15 for any $c \in \mathbb{R}$ there exists a set $N_c \subset \mathbb{R}$ with $\mu(N_c) = 0$ such that

$$\lim_{r\to 0} \fint_{B(x,r)} |f(y) - c| d\mu(y) = \lim_{r\to 0} \fint_{B(x,r)} |f^*(y) - c| d\mu(y) = |f^*(x) - c| \quad \text{for all } x \in \mathbb{R}^n \backslash N_c.$$

Let

$$M := \{x : |f^*(x)| = \infty\}$$

which is a zeroset, since $f^*(x) = f(x)$ μ-a.e. and $f \in L^1_{loc}$.

Let

$$N := M \cup \bigcup_{c\in\mathbb{Q}} N_c.$$

Since $\mathbb{Q}$ is countable, $\mu(N) = 0$. Let $x \in \mathbb{R}^n \setminus N$ and let $c_n \in \mathbb{Q}$ such that $c_n \xrightarrow{n\to\infty} f^*(x)$. Then

$$\fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) \le \fint_{B(x,r)} |f(y) - c_n| d\mu(y) + |c_n - f^*(x)| \xrightarrow{r\to 0^+} |f^*(x) - c_n| + |c_n - f^*(x)|,$$

that is

$$\limsup_{r\to 0^+} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) \le |f^*(x) - c_n| + |c_n - f^*(x)|$$

This holds for any $n \in \mathbb{N}$ and letting $n \to \infty$ we have

$$\limsup_{r\to 0^+} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) = 0,$$

and thus

$$\lim_{r\to 0^+} \fint_{B(x,r)} |f(y) - f^*(x)| d\mu(y) = 0.$$

We can conclude since this holds for any $x \in \mathbb{R}^n \setminus N$, where N satisfies $\mu(N) = 0$. $\square$

Sometimes one wants to work with cubes or ellipsoids shrinking to zero instead of balls. This is possible, as long as they are regular (and μ is doubling). We will restrict here our attention to the Lebesgue measure.

Definition 5.19. We say that a family $\mathcal{F}$ of measurable sets in $\mathbb{R}^n$ is *regular* at $x \in \mathbb{R}^n$ if there is a constant $C > 0$ such that for every $S \in \mathcal{F}$ there is a ball $B(x, r_S)$ such that

$$S \subset B(x, r_S), \quad \mathcal{L}^n(B(x, r_S)) \leq C\mathcal{L}^n(S),$$

and for every $\varepsilon > 0$ there is a set $S \in F$ with $\mathcal{L}^n(S) < \varepsilon$.

Theorem 5.20. *$f \in L^1_{loc}(\mathbb{R}^n, \mathcal{L}^n)$. If $x \in \mathbb{R}^n$ is a Lebesgue point of f and $\mathcal{F}$ is a regular family at $x \in \mathbb{R}^n$, then*

$$\lim_{S\in F, \mathcal{L}^n(S)\to 0} \fint_S f d\mathcal{L}^n = f(x).$$

Proof. For $S \in \mathcal{F}$ let r_S be defined as above. Observe that if $\mathcal{L}^n(S) \to 0$, then $r_S \to 0$. We have

$$\begin{aligned}
\left|\fint_S f d\mathcal{L}^n - f(x)\right| \leq & \fint_S |f(y) - f(x)| d\mathcal{L}^n(y) \\
\leq & \mathcal{L}^n(S)^{-1} \int_{B(x,r_s)} |f(y) - f(x)| d\mathcal{L}^n(y) \\
= & \frac{\mathcal{L}^n(B(x, r_S))}{\mathcal{L}^n(S)} \fint_{B(x,r_S)} |f(y) - f(x)| d\mathcal{L}^n(y) \\
\leq & C \fint_{B(x,r_S)} |f(y) - f(x)| d\mathcal{L}^n(y) \xrightarrow{\mathcal{L}^n(S)\to 0} 0.
\end{aligned}$$

$\square$

Corollary 5.21. *If $f \in L^1_{loc}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$ is a Lebesgue point of f, then for any sequence of cubes Q_i such that $x \in Q_i$,* $\operatorname{diam} Q_i \to 0$ *we have*

$$\lim_{i\to\infty} \fint_{Q_i} f d\mathcal{L}^n = f(x).$$

Corollary 5.22. *Let $F(x) = \int_a^x f(t)dt$ where $f \in L^1(a, b)$. Then $F'(x) = f(x)$ for a.e. $x \in (a, b)$.*

Proof. We have

$$\frac{F(x+h)-F(x)}{h} = \fint_x^{x+h} f(t)dt \xrightarrow{h\to 0} f(x)$$

whenever x is a Lebesgue point of f. $\square$

Exercise 5.23. *Show the following: if $A \subset \mathbb{R}^n$ is a $\mathcal{L}^n$-measurable set, then for almost all $x \in A$*

$$\lim_{r\to 0^+} \frac{|B(x,r)\cap A|}{|B(x,r)|} = 1$$

and for almost all $x \in \mathbb{R}^n \setminus A$

$$\lim_{r\to 0^+} \frac{|B(x,r)\cap A|}{|B(x,r)|} = 0.$$

Hint: $f := \chi_A$

Definition 5.24. Let $A \subset \mathbb{R}^n$ be a measurable set. We say that $x \in \mathbb{R}^n$ is a *density point* of A if

$$\lim_{r\to 0} \frac{|A\cap B(x,r)|}{|B(x,r)|} = 1.$$

Thus the above theorem says that almost every point of a measurable set $A \subset \mathbb{R}^n$ is a density point.

5.4. **Signed (pre-)measures – Hahn decomposition theorem.** We want to apply now and then the Radon-Nikodym theorem to measures which can be positive and negative. For example

$$\mu \llcorner f(A) := \int_A f d\mu \quad A \text{ } \mu\text{-measurable}$$

where $f \in L^1(X, d\mu)$ but f may be positive or negative.

This leads to the so-called *signed measures*. Since e.g. $\mu \llcorner f$ is only defined on measurable sets, we will actually work with pre-measures.

Definition 5.25. Let Σ be a *σ-algebra* and $\mu : \Sigma \to [-\infty, \infty]$. μ is called a *signed premeasure* if

(1) $\mu(\emptyset) = 0$
(2) μ attains at most one of the values $\pm\infty$.
(3) μ is countably additive, i.e. if $(E_k)_{k\in\mathbb{N}}$ are disjoint sets in Σ then

$$\mu(\bigcup E_k) = \sum_{k\in\mathbb{N}} \mu(E_k)$$

Example 5.26.
- If μ_1 and μ_2 are (positive) measures and one of them is finite then $\mu := \mu_1 - \mu_2$ is a signed premeasure on the intersection of the measurable sets of μ_1 and μ_2.

- If ν is a (positive) measure and $f \in L^1(\nu)$ then
$$\mu(E) := \nu \llcorner f = \int_E f d\nu$$
is a signed premeasure on the set of ν-measurable sets. Indeed, if $f \in L^1$ then $f_+, f_- \in L^1$, so
$$\mu(E) = \int_E f^+ d\nu - \int_E f^- d\nu$$
and each of the terms is a (positive) measure.

Theorem 5.27 (Hahn's decomposition theorem)**.** *Let $\mu : \Sigma \to \bar{\mathbb{R}}$ be a signed premeasure on X. Then there are two disjoint sets $X^+, X^- \subset \Sigma$ such that*
$$X = X^+ \cup X^-, \quad X^+ \cap X^- = \emptyset,$$
such that
$$\mu_+ = \mu \llcorner X^+, \quad \mu_- := -\mu \llcorner X_-$$
then $\mu = \mu_+ - \mu_-$ and $\mu_+, \mu_- : \Sigma \to [0, \infty]$ are nonnegative premeasures on X.

Exercise 5.28. *Prove that the decomposition $X = X_+ \cup X-$ is unique up to sets of μ-measure zero, i.e. if $X = \tilde{X}_+ \cup \tilde{X}_-$ is another decomposition, then*
$$\mu(X_+ \setminus \tilde{X}_+) = \mu(\tilde{X}_+ \setminus X_+) = \mu(X_- \setminus \tilde{X}_-) = \mu(\tilde{X}_- \setminus X_-) = 0$$

Example 5.29. Let $\mu := f \llcorner \nu$, for $f \in L^1(\nu)$, i.e.
$$\mu(E) = \int_E f d\nu$$
for ν-measurable f. Then μ is a signed premeasure and
$$\mu(E) = \int_E f^+ d\nu - \int_E f^- d\nu$$
Hence we can take
$$X^+ = \{x : f(x) \geq 0\}$$
$$X^- = \{x : f(x) < 0\}$$
but we can also take
$$\tilde{X}^+ = \{x : f(x) > 0\}$$
$$\tilde{X}^- = \{x : f(x) \leq 0\}$$
In general $X^+ \neq \tilde{X}^+$, and $X^- \neq \tilde{X}^-$. But the sets differ by the set where $f = 0$ and this set has μ-measure zero,
$$\mu(\{x : f(x) = 0\}) = \int_{\{f=0\}} f d\nu = 0.$$

For the proof of Theorem 5.27 we need the notion of positive sets.

Definition 5.30. Let $\mu : \Sigma \to \bar{\mathbb{R}}$ be a signed premeasure. A set $E \in \Sigma$ is called *positive* if

- $\mu(E) > 0$

- $\mu(A) \geq 0$ for any $A \in \Sigma$, $A \subset E$.

Example 5.31. If $E = E_1 \cup E_2$ and $E_1 \cap E_2 = \emptyset$, $\mu(E_1) = 7$, $\mu(E_2) = -3$ then $\mu(E) = 4 > 0$, but E is not positive.

Lemma 5.32. *If $\mu : \Sigma \to \bar{\mathbb{R}}$ is a signed measure then every measurable set E such that $0 < \mu(E) < \infty$ contains a positive set A with $\mu(A) > 0$.*

Proof. If E itself is positive we can take $A = E$.

Otherwise there is $B \subset E$ with $\mu(E) < 0$. Let $n_1 \in \mathbb{N}$ be the smallest positive integer such that there is $B_1 \subset E$ with

$$\mu(B_1) \leq -\frac{1}{n_1}.$$

Observe that $A_1 := E \setminus B_1$ satisfies

Also observe that $\mu(A_1) < \infty$, indeed if $\mu(A_1) = \infty$ then $\mu(B_1) = -\infty$, but then μ attains both $+\infty$ and $-\infty$ which is ruled out. That is $0 < \mu(A_1) < \infty$.

So either A_1 is positive and we take A_1, or we can repeat this construction. Namely in the next step we take $n_2 \in \mathbb{N}$ the smallest positive integer such that there is $B_2 \subset A_1 = E \setminus B_1$ with

$$\mu(B_2) \leq -\frac{1}{n_2}.$$

As above, either $A_2 = E \setminus (B_1 \cup B_2)$ is positive, then we can take $A = A_2$, if not we continue.

If $A_m = E \setminus \bigcup_{j=1}^m B_j$ is positive for some m, we take $A = A_m$ (and have $0 < \mu(A_m) < \infty$), otherwise we obtain an infinite sequence B_j and we set

$$A := E \setminus \bigcup_{j=1}^{\infty} B_j \in \Sigma$$

We claim that $0 < \mu(A) < \infty$ and A is positive.

Firstly, $\mu(E) = \mu(A) + \mu(E \setminus A)$. Since $\mu(E) \in (0, \infty)$ we conclude that if $\mu(A) = +\infty$ then $\mu(E \setminus A) = -\infty$ which is not allowed for a signed measure. So $\mu(A) < \infty$, and we have (observe: $(B_j)_j$ and A are by definition pairwise disjoint!)

$$0 < \mu(E) = \mu(A) + \sum_{j=1}^{\infty} \mu(B_j) \leq \mu(A) - \sum_{j=1}^{\infty} \frac{1}{n_j}.$$

That is

$$\infty > \mu(A) > \sum_{j=1}^{\infty} \frac{1}{n_j} > 0.$$

It remains to prove that A is positive. From the above inequality we conclude that in particular the series on the right-hand side must converge, that is $\lim_{j\to\infty} n_j = \infty$.

If $C \in \Sigma$ and $C \subset A = E \setminus \bigcup_{j=1}^{\infty} B_j$ then for every $m \in \mathbb{N}$ we have $C \subset E \setminus \bigcup_{j=1}^{m} B_j = A_m$. Since n_{m+1} is the smallest positive integer such that there is a set $\tilde{B} \subset A_m$ with $\mu(\tilde{B}) \leq -\frac{1}{n_{m+1}}$ we must have

$$\mu(C) > -\frac{1}{n_{m+1}-1} \xrightarrow{m\to\infty} 0.$$

That is $\mu(C) \geq 0$. Thus A is positive. □

Proof of Theorem 5.27. W.l.o.g. assume that μ does not attain the value $+\infty$, otherwise we work with $-\mu$.

Let

$$M := \sup\left\{\mu(\tilde{A}) : \tilde{A} \in \Sigma,\ \tilde{A} \text{ positive}\right\} \in (-\infty, \infty].$$

Then there is a sequence of positive sets

$$A_1 \subset A_2 \subset A_3 \subset \dots$$

with

$$\mu(A_i) \xrightarrow{i\to\infty} M.$$

Set $A := \bigcup_{i=1}^{\infty} A_i$ then

- A is positive
- $\mu(A) = M < \infty$

It remains to prove that $X \setminus A$ is negative (then we take $X_+ = A$, $X_- = X \setminus A$).

Assume by contradiction that $X \setminus A$ is not negative, there is $E \subset X \setminus A$ with $0 < \mu(E) < \infty$. Hence Lemma 5.32 implies that there is a positive subset $C \subset E$ of positive measure. Now $A \cup C$ is positive and

$$\mu(A \cup C) = \mu(A) + \mu(C) > M$$

which is an obvious contradiction. □

5.5. **Riesz representation theorem.** We will apply now the Radon-Nikodym theorem to classify all continuous linear functionals on $L^p(\mathbb{R}^n)$ for $1 \leq p < \infty$. Let us start first with some general facts about continuous linear mappings between normed spaces.

Theorem 5.33. *Let $L : X \to Y$ be a linear mapping between normed spaces. Then the following conditinos are equivalent*

(1) L is continuous;
(2) L is continuous at 0;
(3) L is bounded, i.e. there is $C > 0$ such that

$$\|Lx\| \leq C\|x\| \quad \text{for all } x \in X \tag{5.6}$$

(4) L is uniformly Lipschitz continuous, i.e. there is $C > 0$ such that $\|Lx - Ly\| \leq C\|x - y\|$ for all $x \in X$.

Remark 5.34. Formally we should use different symbols to denote norms in spaces X and Y. Since it will always be clear in which space we take the norm it is not too dangerous to use the same symbol $\|\cdot\|$ to denote apparently different norms in X and Y.

Proof. The implication (1)⇒(2) is obvious.

(2)⇒(3). Suppose L is continuous at 0 but not bounded. Then there is a sequence $x_n \in X$ such that

$$\|Lx_n\| \geq n\|x_n\|.$$

In particular $\|x_n\| \neq 0$, so

$$0 \leftarrow \left\| L\frac{x_n}{n\|x_n\|} \right\| \geq 1$$

which is a contradiction[25].

(3)⇒(1). Let $x_n \to x$. Then

$$\|Lx - Lx_n\| = \|L(x - x_n)\| \leq C\|x - x_n\| \to 0$$

and hence $Lx_n \to Lx$ which proves continuity of L.

(3) ⇔ (4) is obvious, simply take $y = 0$ (so $Ly = 0$) and use linearity of L. □

Definition 5.35. The number $\|L\| = sup_{\|x\|\leq 1}\|Lx\|$ is called the norm of L. It is often called the *operatornorm*

Exercise 5.36. *Let X and Y be two normed vector spaces. Denote by*

$$L(X, Y) := \{T : X \to Y : \quad T \textit{ linear and continuous}\}$$

the space of linear continouts (aka bounded) maps from X to Y.

For two linear continuous operators $T, S : X \to Y$ and $\lambda, \mu \in \mathbb{R}$ we set

$$(\lambda T + \mu S) : X \to Y$$

defined via

$$(\lambda T + \mu S)\, x := \lambda Tx + \mu Sx.$$

1. *Show that for $T, S \in L(X, Y)$ we also have $\lambda T + \mu S \in L(X, Y)$.*
2. *Show that the operatornorm $\|\cdot\|$ from Lemma 5.37 is indeed a norm on $L(X, Y)$.*

Lemma 5.37. *If $L : X \to Y$ is a linear continuous operator then the norm $\|L\|$ is the smallest number C for which the inequality (5.6) is satisfied.*

That is

$$\|Lx\| \leq \|L\|\,\|x\| \quad \forall x \in X$$

and for each $C \geq 0$ if

$$\|Lx\| \leq C\,\|x\| \quad \forall x \in X$$

[25]Convergence to 0 follows from the continuity of L at 0.

then $C \geq \|L\|$.

Proof. Clearly, if $x = 0$ then

$$\|Lx\| = 0 \leq \|L\|\|0\|.$$

If $x \geq 0$ then

$$\|Lx\| = \|L\frac{x}{\|x\|}\| \, \|x\| \leq \|L\|\|x\|.$$

Now assume that $\|Lx\| \leq C\|x\|$ for all $x \in X$.

If $\|x\| \leq 1$, then $\|Lx\| \leq C\|x\| \leq C$ and hence $\|L\| = sup_{\|x\|\leq 1}\|Lx\| \leq C$. □

Now we want to consider linear bounded *functionals*, which are continuous linear maps $L : X \to \mathbb{R}$.

Example 5.38. If $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p, q < \infty$ are Hölder conjugate, and $g \in L^q(X)$, then

$$L_g(f) := \int_X fgd\mu$$

defines a bounded linear functional on $L^p(\mu)$ and $|Lg| \leq \|g\|_{L^q}$. Indeed, Hölder's inequality yields

$$|L_g f| = \left|\int_X fgd\mu\right| \leq \|g\|_{L^q}\|f\|_{L^p}.$$

Similarly if $g \in L^\infty(\mu)$, then

$$L_g f = \int_X fgd\mu$$

defines a bounded linear functional on $L^1(\mu)$ and $\|L^g\| \leq \|g\|_{L^\infty}$.

The following theorem states that these are the only linear functionals on $L^p(\mathbb{R}^n)$ as long as $1 \leq p<\infty$ (similar statement hold for reasonable μ).

Theorem 5.39 (Riesz representation theorem)**.** *Let* $1 \leq p < \infty$ *and* $L : L^p(\mathbb{R}^n) \to \mathbb{R}$ *be a bounded linear functional.*

$$|L[f]| \leq \|T\|\|f\|_{L^p(\mathbb{R}^n)} \tag{5.7}$$

Denoting by $p' = \frac{p}{p-1}$ *the Hölder conjugate there exists* $g \in L^{p'}(\mathbb{R}^n)$ *such that*

$$L[f] = \int_{\mathbb{R}^n} fg \quad \forall f \in L^p(\mathbb{R}^n).$$

and

$$\|g\|_{L^{p'}(\mathbb{R}^n)} = \|T\|. \tag{5.8}$$

The function g *is unique in the sense that if there are* g_1 *and* g_2 *satisfying the above then* $g_1 = g_2$ *a.e..*

Remark 5.40.
- The set of bounded linear functionals on $L^p(\mathbb{R}^n)$ is usually denoted by $(L^p(\mathbb{R}^n))^*$. It is a vector space (we will discuss more properties in Functional Analysis) and called the *dual space* to $L^p(\mathbb{R}^n)$. What Theorem 5.39 says is that there exists an isometric isomorphism from the space of bounded linear functionals L on $L^p(\mathbb{R}^n)$ with $L^{p'}(\mathbb{R}^n)$, i.e.
$$(L^p(\mathbb{R}^n))^* \cong L^{p'}(\mathbb{R}^n).$$
Often one says that *the dual space to $L^p(\mathbb{R}^n)$, i.e. $(L^p(\mathbb{R}^n))^*$, is $L^{p'}(\mathbb{R}^n)$*. What is important here (but often this is not so explicitely clear for other spaces) is the *pairing*, i.e. in which sense this "identity" holds. We observe this is the L^2-scalar products (even if L^p-functions do not in general belong to L^2. Any element in $L \in (L^p(\mathbb{R}^n))^*$ can be identified with some $g \in L^{p'}(\mathbb{R}^n)$ via the *L^2-pairing*
$$L(f) = \langle f, g\rangle.$$

Proof of Theorem 5.39. Uniqueness Assume that there are g_1 and g_2 that both satisfy the claims, then
$$\int (g_1 - g_2) f = 0 \quad \forall f \in L^p(\mathbb{R}^n).$$
In particular
$$\int (g_1 - g_2)\varphi = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$
By the fundamental theorem of the calculus of variations, Lemma 4.45 we have $g_1 = g_2$ a.e.

Existence Let $K \subset \mathbb{R}^n$ be compact with $\mathcal{L}^n(\partial K) = 0$ and consider for $\mathcal{L}^n$-measurable $A \subset \mathbb{R}^n$
$$\nu(A) := L[\chi_{A\cap K}].$$
Observe that $|\nu(A)| \leq \|\chi_K\|_{L^p(\mathbb{R}^n)} < \infty$.

We claim that ν is a *signed* pre-measure on the σ-Algebra of $\mathcal{L}^n$-measurable sets. For this let $(A_k)_{k=1}^\infty$ be disjoint $\mathcal{L}^n$-measurable sets and set $A := \bigcup_k A_k$. Then for all $x \in \mathbb{R}^n$ we have
$$f(x) := \chi_{A\cap K}(x) = \lim_{k\to\infty} f_k(x)$$
where $f_k := \sum_{i=1}^k \chi_{A_i \cap K}$. Since K is compact, $f \in L^p(\mathbb{R}^n)$. Moreover, since the $(A_i)_i$ are pairwise disjoint, so by Lebesgue monotone convergence, Corollary 3.8,
$$\|f_k\|_{L^p(\mathbb{R}^n)}^p = \|\sum_{i=1}^k \chi_{A_i\cap K}\|_{L^1(\mathbb{R}^n)} \xrightarrow{k\to\infty} \|f\|_{L^1(\mathbb{R}^n)} = \|f\|^p_{L^p(\mathbb{R}^n)}.$$
Sine $p < \infty$[26], in view of Theorem 3.52, we then have
$$\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0 \tag{5.9}$$

[26]here we use this crucially

and thus

$$\nu(\bigcup_{i=1}^{\infty} A_i) = L(f) = \lim_{k\to\infty} L(f_k) = \sum_{i=1}^{\infty} \nu(A_i).$$

That is ν is a signed premeasure.

By Hahn decomposition theorem, Theorem 5.27, we can split ν into two pre-measures ν_+ and ν_-,

$$\nu = \nu_+ - \nu_-.$$

Then ν_+, ν_- both extend to a Radon measure, which by (5.7) satisfies $\nu_\pm \ll \mathcal{L}^n$. Indeed, since ν is a signed premeasure and $|\nu(\mathbb{R}^n)| < \infty$ we have that $\nu_+(\mathbb{R}^n), \nu_-(\mathbb{R}^n) < \infty$ (because otherwise they would both be infinite which would contradict $|\nu(\mathbb{R}^n)| < \infty$).

By Radon-Nikodym theorem for $g_K := \chi_K D_{\mathcal{L}^n}\nu_+ - \chi_K D_{\mathcal{L}^n}\nu_-$ which is $\mathcal{L}^n$-a.e. finite and $\mathcal{L}^n$-integrable (because each ν_+ and ν_- are finite measures), we have

$$\nu(A) = \nu(A\cap K) = \int_{\mathbb{R}^n} \chi_A g_K(x)\, d\mathcal{L}^n(x).$$

In particular we have $g_K(x) \in L^1(\mathbb{R}^N)$, and for all $A \subset \mathbb{R}^n$ $\mathcal{L}^n$-measurable

$$L(\chi_{A\cap K}) = \int_{\mathbb{R}^n} \chi_A g_K(x)\, d\mathcal{L}^n(x).$$

Now let $f \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$, $f \geq 0$ *a.e.* in $\mathbb{R}^n$ then there exists an approximation

$$f_k(x) = \sum_{i=1}^{\infty} a_k \chi_{A_k}(x)$$

such that $f_k \to f$ a.e. in $\mathbb{R}^n$. In particular

$$|f(x)|^p = \lim_{k\to\infty} |f_k(x)|^p \quad \text{a.e. in } \mathbb{R}^n.$$

By the monotone convergence theorem we conclude that $\lim_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} = \|f\|_{L^p(\mathbb{R}^n)} < \infty$. Here we used Exercise 3.18. Thus from Theorem 3.52 we find that $\|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$.

We then have

$$L(\chi_K f) = L\left(\chi_K\left(f - f_k\right)\right) + L\left(\chi_K f_k\right) = L(\chi_K(f-f_k)) + \int_{\mathbb{R}^n} f_k g_K\ d\mathcal{L}^n.$$

Since L is uniformly Lipschitz and $f - f_k \xrightarrow{k\to\infty} 0$ in $L^p(\mathbb{R}^n)$ we have $L(\chi_K(f-f_k)) \xrightarrow{k\to\infty} 0$.

Also, by the monotone convergence theorem

$$\begin{aligned}\lim_{k\to\infty} \int_{\mathbb{R}^n} f_k g_K\ d\mathcal{L}^n &= \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k (g_K)_+\ d\mathcal{L}^n - \lim_{k\to\infty} \int_{\mathbb{R}^n} f_k (g_K)_-\ d\mathcal{L}^n \\ &= \int_{\mathbb{R}^n} f(g_K)_+\ d\mathcal{L}^n + \int_{\mathbb{R}^n} f(g_K)_+\ d\mathcal{L}^n = \int_{\mathbb{R}^n} f g_K\ d\mathcal{L}^n.\end{aligned}$$

In the last step we have used that $\|f g_K\|_{L^1(\mathbb{R}^n)} \leq \|f\|_{L^\infty(\mathbb{R}^n)} \|g_K\|_{L^1(\mathbb{R}^n)} < \infty$.

Thus, we have shown

$$L(\chi_K f) = \int_{\mathbb{R}^n} f g_K \, d\mathcal{L}^n \quad \forall f \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n), \ f \geq 0.$$

Splitting $f = f_+ + f_-$ we conclude that we don't need the sign condition on f.

$$L(\chi_K f) = \int_{\mathbb{R}^n} f g_K \, d\mathcal{L}^n \quad \forall f \in L^\infty(\mathbb{R}^n) \cap L^1(\mathbb{R}^n).$$

For now we have $g_K \in L^1(\mathbb{R}^n)$ only. Next we need to show $\|g_K\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|$.

If $p = 1$ and thus $p' = \infty$ let $x \in \mathbb{R}^n \backslash \partial K$ be any Lebesgue point of g_K. If we choose $f := \chi_{B(x,r)}$ for $r < \operatorname{dist}(x, \partial K)$

$$\left| \int_{B(x,r)} g_K \, d\mathcal{L}^n \right| = \left| L\chi_{K\cap B(x,r)} \right| \leq \|T\| \mathcal{L}^n(K \cap B(x,r)).$$

By the assumptions on r if $x \in K$ then $\mathcal{L}^n(K \cap B(x,r)) = \mathcal{L}^n(B(x,r))$. If $x \notin K$ then $\mathcal{L}^n(K \cap B(x,r)) = 0$. In either case by the Lebesgue differentiation theorem we can let $r \to 0$ and have

$$|g_K(x)| \leq \|T\|$$

which holds for any Lebesgue point $x \in \mathbb{R}^n \backslash \partial K$. Since $\mathcal{L}^n(\partial K) = 0$ we have $g_k(x) \leq \|T\|$ for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$, and thus

$$\|g_K\|_{L^\infty} \leq \|T\|.$$

Now assume $p \in (1, \infty)$. We argue *by duality*.

Formally, this is easy: For $p' = \frac{p}{p-1} \in (1, \infty)$ we have

$$\|g_K\|_{L^{p'}}^{p'} = \int_{\mathbb{R}^n} g_K \, g_K |g_K|^{p'-2} = T[|g_K|^{p'-2} g_K] \leq \|T\| \||g_K|^{p'-2} g_K\|_{L^p} = \|T\| \|g_K\|_{L^{p(p'-1)}}^{p'-1}$$

so that (since $p(p'-1) = p'$, we can divide on both sides $\|g_K\|_{L^{p(p'-1)}}^{p'-1} \equiv \|g_K\|_{L^{p'}}^{p'-1})$) we have found

$$\|g_K\|_{L^{p'}} \leq \|T\|.$$

Essentially that will be our argument, but the above only works if we already know that $|g_K\|_{L^{p'}}^{p'-1}) < \infty$ (which is pretty much what we want to show). The above type of argument is called an *a priori estimate*. To make this precise we need mollification (a technique which is often sweeped under the rug by saying *by approximation* or *by density*):

Let $f \in L^1(\mathbb{R}^n)$ with $f = 0$ a.e. outside a compact set $\tilde{K}$. For $\varepsilon > 0$ we denote by f_ε the usual convolution $f_\varepsilon = f * \varepsilon^{-n}\eta(\cdot/\varepsilon)$ (with symmetric kernel η). Then by Exercise 4.21

$$L(\chi_K f_\varepsilon) = \int_{\mathbb{R}^n} f_\varepsilon \, g_K \, d\mathcal{L}^n = \int_{\mathbb{R}^n} f (g_K)_\varepsilon \, d\mathcal{L}^n.$$

Observe that since $\operatorname{supp} g_K \subset K$ we have that $(g_K)_\varepsilon \in C_c^\infty(\mathbb{R}^n)$.

Then we may pick (here we use that $p \in (1,\infty)$)

$$f(x) := \begin{cases} |(g_K)_\varepsilon(x)|^{\frac{p}{p-1}-1} & \text{if } (g_K)_\varepsilon(x) \geq 0 \\ -\,|(g_K)_\varepsilon(x)|^{\frac{p}{p-1}-1} & \text{if } (g_K)_\varepsilon(x) < 0 \end{cases} \tag{5.10}$$

(you can check this is measurable). Then we have $f \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$. Indeed, this follows from (observe: $\frac{p}{p-1} - 1 > 0$)

$$|f(x)| \leq |(g_K)_\varepsilon(x)|^{\frac{p}{p-1}-1} \leq \|(g_K)_\varepsilon(x)\|_{L^\infty}^{\frac{p}{p-1}-1},$$

so $f \in L^\infty(\mathbb{R}^n)$, but also f has compact support. Thus we have

$$\begin{aligned}\int_{\mathbb{R}^n} |(g_K)_\varepsilon|^{p'} =& \int_{\mathbb{R}^n} f(x)\,(g_K)_\varepsilon \\ =& L(\chi_K f) \leq \|T\| \underbrace{\|\,((g_K)_\varepsilon)^{\frac{p}{p-1}-1}\,\|_{L^p(\mathbb{R}^n)}}_{<\infty}.\end{aligned}$$

A short computation yields

$$(p'-1)p = \frac{p}{p-1} = p'$$

and thus

$$\|\,((g_K)_\varepsilon)^{\frac{p}{p-1}-1}\,\|_{L^p(\mathbb{R}^n)} = \|(g_K)_\varepsilon\|_{L^{p'}(\mathbb{R}^n)}^{\frac{p'}{p}}$$

We conclude that

$$\|(g_K)_\varepsilon\|_{L^{p'}(\mathbb{R}^n)}^{p'} = \int_{\mathbb{R}^n} |(g_K)_\varepsilon|^{p'} \leq \|T\| \|(g_K)_\varepsilon\|_{L^p(\mathbb{R}^n)}^{\frac{p'}{p}}$$

Observing that $p' - \frac{p'}{p} = p'(1-\frac{1}{p}) = \frac{p'}{p'} = 1$ we have shown that

$$\|(g_K)_\varepsilon\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|.$$

This holds for all $\varepsilon > 0$. Since $g_K \in L^1(\mathbb{R}^N)$ we have that $(g_K)_\varepsilon$ converges to g_K in L^1 as $\varepsilon \to 0$, which means for some $\varepsilon_i \to 0$ the functions $(g_K)_{\varepsilon_i}$ converge a.e. to g_K. From Exercise 3.14 we finally obtain

$$\|g_K\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|.$$

Removing the K. Now from the construction of g_K we see that if $K \subset K'$ both compact with ∂K and $\partial K'$ both zero sets, then $g_K = g_{K'}$ a.e. in K and $|g_K| \leq |g_{K'}|$ a.e. in $\mathbb{R}^n$.

So $g(x) := g_K(x)$ where K is a large compact set containing x makes g well defined and we have

$$\|\chi_K g\|_{L^{p'}(\mathbb{R}^N)} = \|g_K\|_{L^{p'}(\mathbb{R}^N)} \leq \|T\| \quad \forall K \subset \mathbb{R}^n \text{ compact}.$$

By monotonce convergence theorem we have

$$\|g\|_{L^{p'}(\mathbb{R}^n)} = \|\,|g|\,\|_{L^{p'}(\mathbb{R}^n)} = \lim_{k\to\infty} \|\chi_{B(0,k)}\,|g|\,\|_{L^{p'}(\mathbb{R}^n)} \leq \|T\|$$

and we have

$$L(f) = \int_{\mathbb{R}^n} gf \; d\mathcal{L}^n(\mathbb{R}^n) \quad \forall f \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n) : \quad \operatorname{supp} f \text{ compact}.$$

(simply take K to contain $\operatorname{supp} f$).

It remains to show that

$$L(f) = \int_{\mathbb{R}^n} gf \quad \forall f \in L^p(\mathbb{R}^n). \tag{5.11}$$

We do this *by density* (here we use that $p < \infty$). Any $f \in L^p(\mathbb{R}^n)$ can be approximated by $f_k \in C_c^\infty(\mathbb{R}^n)$. Then we have

$$\begin{aligned} L(f) =& L(f - f_k) + L(f_k) = L(f - f_k) + \int_{\mathbb{R}^n} gf_k \\ =& L(f - f_k) + \int_{\mathbb{R}^n} gf + \int_{\mathbb{R}^n} g(f_k - f). \end{aligned}$$

Now by boundedness of L

$$|L(f - f_k)| \leq \|T\| \|f - f_k\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

Moreover by Hölder's inequality

$$\left| \int_{\mathbb{R}^n} g(f_k - f) \right| \leq \|g\|_{L^p(\mathbb{R}^n)} \|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

So we have established (5.11).

The last thing to establish is (5.8). We have already

$$\|g\|_{L^{p'}} \leq \|T\|.$$

On the other hand by Hölder's inequality

$$|T(f)| = \left| \int fg \right| \leq \|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^{p'}(\mathbb{R}^n)} \quad \forall f \in L^p(\mathbb{R}^n).$$

In view of Lemma 5.37 we find $\|g\| \leq \|T\|$. □

Remark 5.41. Theorem 5.39 is not true for $p = \infty$, since (5.9) may fail for $p = \infty$ (cf. Exercise 3.53). Indeed any linear functional L on $L^\infty(\mathbb{R}^n)$ can be written as an integration $L(f) = \int f d\mu$ where μ is a *finite* additive signed measures, which are absolutely continuous with respect to $\mathcal{L}^n$. The Radon-Nikodym theorem does not apply for *finite* additive measures, but for *infinitely* additive measures.

See also Lemma 11.12.

Exercise 5.42. *Use Theorem 5.39 and Exercise 4.34 to prove the following slight (but useful!) generalization of Riesz representation theorem.*

Assume $1 \leq p \leq \infty$ *and assume* $L : C_c^\infty(\mathbb{R}^n) \to \mathbb{R}$ *is a linear map that satisfies for some* $\Lambda \geq 0$

$$|L\varphi| \leq \Lambda \|\varphi\|_{L^p(\mathbb{R}^n)} \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

Denoting by $p' = \frac{p}{p-1}$ *the Hölder conjugate there exists* $g \in L^{p'}(\mathbb{R}^n)$ *such that*

$$L[f] = \int_{\mathbb{R}^n} fg \quad \forall f \in C_c^\infty(\mathbb{R}^n).$$

and

$$\|g\|_{L^{p'}(\mathbb{R}^n)} = \|T\|.$$

The function g *is unique in the sense that if there are* g_1 *and* g_2 *satisfying the above then* $g_1 = g_2$ *a.e..*

Hint: Show that the following functional $\tilde{L} : L^p(\mathbb{R}^n) \to \mathbb{R}$ *is well-defined, linear, and bounded. For* $f \in L^p(\mathbb{R}^n)$ *take any approximation* $f_k \in C_c^\infty(\mathbb{R}^n)$, $f_k \xrightarrow{k\to\infty} f$ *in* $L^p(\mathbb{R}^n)$. *Then we define*

$$\tilde{L}f := \lim_{k\to\infty} Lf_k.$$

(You have to check that $\tilde{L}(f) \in \mathbb{R}$ *and that this value is independent of the choice of approximation!)*

Exercise 5.43 (Minkowski's integral inequality)**.** *Let* $f \in C_c^\infty(\mathbb{R}^{k+\ell})$ *and assume that* $1 < p < \infty$. *We denote variables in* $\mathbb{R}^{k+\ell}$ *by* (x,y), $x \in \mathbb{R}^k$, $y \in \mathbb{R}^\ell$. *Show the following*

$$\left(\int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^\ell} |f(x,y)| dy\right)^p dx\right)^{\frac{1}{p}} \leq \int_{\mathbb{R}^\ell} \left(\int_{\mathbb{R}^k} |f(x,y)|^p dx\right)^{\frac{1}{p}} dy$$

Use the duality characterization of L^p *as in* (5.10), *and Fubini's theorem.*

Let us also record (without proof) the following version of the Riesz representation theorem that classify the set of Radon measures wiht linear functions on

$$C_c(\mathbb{R}^n) := \{f \in C^0(\mathbb{R}^n) : \quad \operatorname{supp} f \text{ bounded}\}.$$

Theorem 5.44 (Riesz Representation Theorem)**.** *Let*

$$L : C_c(\mathbb{R}^n; \mathbb{R}^m) \to \mathbb{R}$$

be a linear functional satisfying

$$\sup \{L(f) : \quad f \in C_c(\mathbb{R}^n; \mathbb{R}^m),\ |f| \leq 1, \operatorname{supp} f \subset K\} < \infty$$

for every compact set $K \subset \mathbb{R}^n$.

Then there exists a Radon measure μ *on* $\mathbb{R}^n$ *and a* μ*-measurable function* $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ *such that*

$$|\sigma(x)| = 1 \quad \mu\text{-a.e. } x$$

and

$$L(f) = \int_{\mathbb{R}^n} f \cdot \sigma d\mu \quad \forall f \in C_c(\mathbb{R}^n, \mathbb{R}^m).$$

For a proof see Cf. [Evans and Gariepy, 2015, Theorem 1.38]

6. Transformation Rule

In this section we discuss the transformation rule, and will assume always $\mu = \mathcal{L}^n$.

The change of variables formula is a general version of the substitution rule. If $\Phi' \neq 0$ we have

$$\int_{[a,b]} f(\Phi(x))|\Phi'(x)|dx = \int_{\Phi([a,b])} f(x)dx.$$

Why do we write $|\Phi'(x)|$ not $\Phi'(x)$ (as we learned in Calculus)?

Assume that $\Phi(x) = -x$ then for $a < b$ (i.e. $-a > -b$),

$$\int_{[a,b]} f(-x)dx = \int_a^b f(-x)dx = -\int_{-a}^{-b} f(x)dx = \int_{-b}^{-a} f(x)dx = \int_{[-b,-a]} f(x) \tag{6.1}$$

I.e. *orientation* is accounted for differently with the notation $\int_a^b$ than with $\int_{[a,b]}$.

The one-dimensional $\Phi'(x)$ becomes $\det(D\Phi(x))$ because (as we have also seen in Theorem 1.81), this is the change of volume of the unit cube under the linear map $D\Phi(x)$, cf. Figure 6.1 and Figure 1.6.

Definition 6.1. Let $X, Y \subset \mathbb{R}^n$ be open.

$\Phi : X \to Y$ is a C^k-diffeomorphism if

- $\Phi : X \to Y$ is bijective
- Φ, Φ^{-1} are C^k

One can show that this implies

$$\det(D\Phi(x)) \neq 0 \quad \text{in } X.$$

Theorem 6.2 (Change of variables formula/transformation rule)**.** *Let $\Phi : X \to \Phi(X)$ be a C^1-diffeomorphism between open and bounded sets X, $\Phi(X) \subset \mathbb{R}^n$.*

(1) for any measurable $\Omega \subset X$ we have

$$\mathcal{L}^n(\Phi(\Omega)) = \int_\Omega |\det(D\Phi(x))|dx.$$

(2) $f : Y \to \bar{R}$ is integrable in Y if and only if $f \circ \Phi |\det(D\Phi)|$ is integrable in X and if that is the case,

$$\int_Y f dx = \int_X f \circ \Phi |\det(D\Phi(x))|dx.$$

or equivalently

$$\int_{\Phi(X)} f dx = \int_X f \circ \Phi |\det(D\Phi(x))|dx.$$

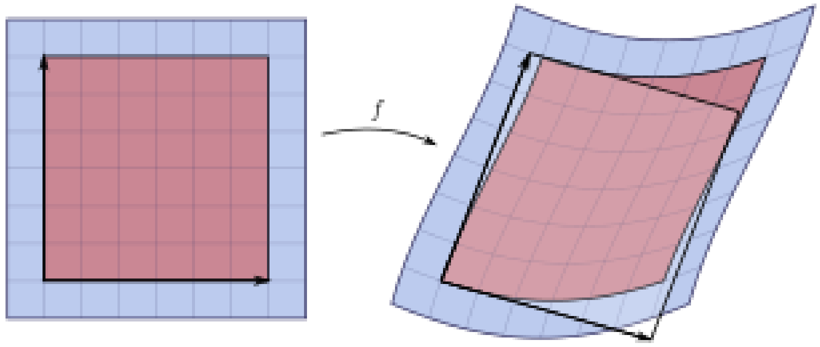


FIGURE 6.1. A nonlinear map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ sends a small square (left, in red) to a distorted parallelogram (right, in red). The Jacobian at a point gives the best linear approximation of the distorted parallelogram near that point (right, in translucent white), and the Jacobian determinant gives the ratio of the area of the approximating parallelogram to that of the original square. By Blacklemon67 - Own work, CC BY-SA 3.0, wikipedia

Lemma 6.3. *Under the assumptions of Theorem 6.2 for any $\Omega \subset X$ measurable*

$$\mathcal{L}^n(\Phi(\Omega)) = \int_\Omega |\det(D\Phi(x))| dx.$$

Proof. Set

$$\mu(A) := \mathcal{L}^n(\Phi(A \cap X)) \quad A \text{ is } \mathcal{L}^n\text{-measurable}$$

Since Φ is a C^1-diffeomorphism this induces a Radon measure and $\mu \ll \mathcal{L}^n$, cf. Exercise 1.80, Exercise 1.13 and Corollary 1.78.

By Radon-Nikdoym theorem, Theorem 5.13, we have

$$\mu(A) = \int_A D_\mu \mathcal{L}^n \, d\mathcal{L}^n.$$

and

$$D_\mu \mathcal{L}^n(x) = \lim_{r \downarrow 0} \frac{\mu(B(x,r))}{\mathcal{L}^n(B(x,r))} \quad \text{for } \mathcal{L}^n - a.e.\ x.$$

So fix $x \in X$ (by exhaustion argument using the monotone convergence theorem we may assume $\operatorname{dist}(x, \partial X) > 0$) and $r \ll 1$. We then have from Taylor's theorem for each $z \in B(x,r)$, since $\Phi \in C^1$

$$|\Phi(z) - \Phi(x) - D\Phi(x)(z-x)| \le o(r).$$

Thus, since $D\Phi(x)$ is invertible matrix,

$$\Phi(B(x,r)) \subset \Phi(x) + \underbrace{D\Phi(x)}_{\in \mathbb{R}^{n\times n}}(B(x,r)-x) + B_{o(r)}(0) = \Phi(x) + D\Phi(x)\left((B(x,r)-x) + D\Phi(x)^{-1} B_{o(r)}(0)\right)$$

Consequently by translation invariance of the Lebesgue measure and Theorem 1.81.

$$\begin{aligned}\mu(B(x,r)) =&\mathcal{L}^n(\Phi(B(x,r))) \leq \mathcal{L}^n\left(D\Phi(x)\left((B(x,r)-x)+D\Phi(x)^{-1}B_{o(r)}(0)\right)\right)\\ =&|\det(D\Phi(x))|\,\mathcal{L}^n\left((B(x,r)-x)+D\Phi(x)^{-1}B_{o(r)}(0)\right)\\ =&|\det(D\Phi(x))|\,\mathcal{L}^n\left(B(x,r)+D\Phi(x)^{-1}B_{o(r)}(0)\right)\\ \leq&|\det(D\Phi(x))|\mathcal{L}^n(B(x,r+o(r))).\end{aligned}$$

Thus

$$\frac{\mu(B(x,r))}{\mathcal{L}^n(B(x,r))} \leq |\det(D\Phi(x))|\frac{\mathcal{L}^n(B(x,r+o(r)))}{\mathcal{L}^n(B(x,r))}.$$

Now $\mathcal{L}^n(B(x,R)) = c_nR^n$, so we have

$$\frac{\mu(B(x,r))}{\mathcal{L}^n(B(x,r))} = |\det(D\Phi(x))|\left(\frac{r+o(r)}{r}\right)^n \xrightarrow{r\to 0^+} |\det(D\Phi(x))|.$$

So

$$D_\mu\mathcal{L}^n(x) = |\det(D\Phi(x))| \quad \text{for } \mathcal{L}^n\text{-a.e. } x$$

and we can conclude. □

Lemma 6.4. *Under the assumptions of Theorem 6.2 assume that $f : Y \to \bar{R}$ is measurable. Then*

$$\int_Y f d\mathcal{L}^n = \int_X f\circ\Phi|\det(D\Phi(x))|dx.$$

Proof. First we assume $f \geq 0$. Since Φ is a diffeomorphism (and in particular maps open sets into open sets) $f\circ\Phi$ is $\mathcal{L}^n$-measurable, and so is $f\circ\Phi|\det D\Phi|$.

We approximate f by simple functions

$$f_k(x) := \sum_{i=1}^k a_i\chi_{A_i}(x)$$

wher $a_i \geq 0$ and $A_i \subset Y$ are measurable, and

$$f_k(x) \to f(x) \quad \text{for every } x\in Y\text{, and monotonely.}$$

Set $B_i := \Phi^{-1}(A_i)$ (which is measurable) and $B_i \subset X$.

Then

$$\begin{aligned}f_k\circ\Phi(x) =&\sum_{i=1}^k a_i\chi_{A_i}\circ\Phi(x)\\ =&\sum_{i=1}^k a_i\chi_{B_i}(x)\end{aligned}$$

So with Lemma 6.3,

$$\begin{aligned}\int_Y f_k dx &= \sum_{i=1}^k a_i \mathcal{L}^n(A_i)\\ &= \sum_{i=1}^k a_i \mathcal{L}^n(\Phi(B_i))\\ &= \sum_{i=1}^k a_i \int_{B_i} |\det(D\Phi)|\\ &= \int_X f_k \circ \Phi \int_{B_i} |\det(D\Phi)|\end{aligned}$$

We conclude the case $f \geq 0$ by using monotone convergence theorem.

For general f we set $f = f_+ - f_-$ and use twice the above argument. □

We will not go into applications here (the polar coordinates and related coordinate changes one being the most prominent ones and we did those in Advanced Calculus).

6.1. **Area formula and integration on manifolds.** Let $\mathcal{M} \subset \mathbb{R}^N$ be a (subset of an) n-dimensional manifold. For $f : \mathcal{M} \to \mathbb{R}$ we can then define

$$\int_{\mathcal{M}} f d\mathcal{H}^n := \int_{\mathbb{R}^N} f d\left(\mathcal{M} \llcorner \mathcal{H}^n\right),$$

because f is defined on $(\mathcal{M} \llcorner \mathcal{H}^n)$-a.e. point. In Advanced Calculus I we talked about another way of defining the integral on $\mathcal{M}$:

Assume that there is $\theta_0 \in \mathcal{M}$ and an open neighborhood $U \subset \mathbb{R}^n$ such that $\Phi : U \to \mathcal{M}$ is a parametrization of a neighborhood $\theta_0 \in \partial\Omega$ then we defined

$$\int_{\mathcal{M} \cap \Phi(U)} f(\theta) d\theta := \int_U f(\Phi(z)) |\mathrm{Jac}(\Phi)(z)| \, dz.$$

Here we defined the *Jacobian* $|\mathrm{Jac}(\Phi)(z)|$ as follows

$$|\mathrm{Jac}(\Phi)(z)| := \sqrt{\left|\det_{\mathbb{R}^{n\times n}} (\underbrace{D\Phi^t(z)}_{\mathbb{R}^{n\times N}} \underbrace{D\Phi(z)}_{\mathbb{R}^{N\times n}})\right|} \equiv \sqrt{\left|\det_{\mathbb{R}^{n\times n}} \left((\langle \partial_\alpha \Phi(z), \partial_\beta \Phi(z)\rangle_{\mathbb{R}^N})_{\alpha,\beta=1,\ldots,n}\right)\right|}$$

We want to show that this coincides with the Hausdorff-measure integral. This is usually referred to as the *area formula*.

Theorem 6.5 (Area formula). *Let $\Phi \in C^1(\mathbb{R}^n, \mathbb{R}^N)$ where $n \leq N$. For each $f \in L^1(\mathbb{R}^n)$ and each $\mathcal{L}^n$-measurable subset $A \subset \mathbb{R}^n$ such that $\Phi : A \to \Phi(A)$ is one-to-one*

$$\int_{\Phi(\mathbb{R}^n) \subset \mathbb{R}^N} f \circ \Phi^{-1}(y) d\mathcal{H}^n(y) := \int_{\mathbb{R}^n} f(z) |\mathrm{Jac}(\Phi)(z)| \, d\mathcal{L}^n(z).$$

Remark 6.6.

- There exist generalizations for Φ which are not one-to-one.

- Having $f : \mathbb{R}^n \to \mathbb{R}$ instead of $f : \mathcal{M} \to \mathbb{R}$ avoids dealing with integrability. But this can be done with density arguments

Proof. By the usual approximation arguments (see the transformation rule) it suffices to prove the statement for $f = \chi_A$, so we need to prove for $\mathcal{L}^n$-measurable sets A,

$$\mathcal{H}^n(\Phi(A)) = \int_A |\mathrm{Jac}(\Phi)(z)| \, d\mathcal{L}^n(z).$$

Observe that since Φ is Lipschitz and $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$ we have that

$$\nu(A) := \mathcal{H}^n(\Phi(A)) \quad A \subset \mathbb{R}^n \text{ is } \mathcal{L}^n\text{-measurable}$$

extends to a Radon measure in $\mathbb{R}^n$. By Radon-Nikodyn we then have

$$\mathcal{H}^n(\Phi(A)) = \nu(A) = \int_A \frac{d\nu}{d\mathcal{L}^n}(z) d\mathcal{L}^n(z).$$

where for μ-a.e. $z \in \mathbb{R}^n$,

$$\frac{d\nu}{d\mathcal{L}^n}(z) = \lim_{r \to 0^+} \frac{\mathcal{H}^n(\Phi(B(z,r)))}{\mathcal{L}^n(B(z,r))}.$$

As in the argument for the transformation rule, since Φ is smooth, $\mathcal{H}^n$ is translation invariant, the claim follows by Tailor's theorem once we can show

$$\mathcal{H}^n(D\Phi(x)A) = \mathrm{Jac}(\Phi)(x)\mathcal{L}^n(A).$$

This is the content of the following Lemma 6.7. $\square$

Lemma 6.7. *Let $L : \mathbb{R}^n \to \mathbb{R}^N$ be a linear map, i.e. $L \in \mathbb{R}^{N \times n}$. Then*

$$\mathcal{H}^n(L(A)) = \sqrt{\det(L^t L)} \mathcal{L}^n(A).$$

Proof. By polar decomposition (cf. [Evans and Gariepy, 2015, Theorem 3.5] there exists a matrix $O \in \mathbb{R}^{N \times n}$ and $S \in \mathbb{R}^{n \times n}$ such that $O \in O(n, N)$, i.e. $O^t O = I_{n \times n}$, i.e.

$$\langle Ox, Oy \rangle_{\mathbb{R}^N} = \langle x, y \rangle_{\mathbb{R}^n} \quad \forall x, y \in \mathbb{R}^n,$$

and

$$S^t = S,$$

and we have

$$L = OS.$$

We then have

$$\det(L^t L) = \det(S^t O^t O S) = \det(S^t S) = (\det(S))^2,$$

that is

$$\sqrt{\det(L^t L)} = |\det(S)|.$$

The conditions on O simply mean that

$$O = (o_1, \ldots, o_n) \in \mathbb{R}^{N \times n}$$

for orthonormal vectors $o_i \in \mathbb{R}^N$, $i = 1, \ldots, n$. We can extend $(o_i)_{i=1}^n$ to a orthonormal basis $(o_i)_{i=1}^N$ of $\mathbb{R}^N$. Set

$$\tilde{O} := (o_1, \ldots, o_n, o_{n+1}, \ldots, o_N) \in \mathbb{R}^{N \times N}.$$

Then $\tilde{O} \in O(n)$, i.e. $\tilde{O}^t \tilde{O} = I_{N \times N}$. By Exercise 1.16

$$\mathcal{H}^n(LA) = \mathcal{H}^n(\tilde{O}^t OSA).$$

Now

$$\tilde{O}^t O = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & \ldots & 0 \\ \vdots & & \vdots & \\ 0 & 0 & \ldots & 0 \end{pmatrix} = \begin{pmatrix} I_{n \times n} \\ 0_{N-n \times n} \end{pmatrix} \in \mathbb{R}^{N \times n}$$

In particular $\tilde{O}^t LA \subset \mathbb{R}^n \times \{0\}$. So again by Exercise 1.16

$$\mathcal{H}^n(LA) = \mathcal{H}^n(\underbrace{SA}_{\subset \mathbb{R}^n}).$$

Since $\mathcal{H}^n = \mathcal{L}^n$ in $\mathbb{R}^n$ we have by the transformation rule on $\mathbb{R}^n$,

$$\mathcal{H}^n(LA) = \mathcal{L}^n(SA) = |\det(S)|\mathcal{L}^n(A) = \sqrt{\det(L^t L)}\mathcal{L}^n(A).$$

□

Part 2. Analysis II: L^p & Sobolev spaces (with sprinkles of Functional Analysis)

For the Functional Analysis parts we use texts from [Brezis, 2011], lecture notes by Michael Struwe (Funktionalanalysis; in German), and Hajłasz (Functional Analysis), as well as [Clason, 2020]. An excellent (readable) source for Fourier Analysis is [Grafakos, 2014]. For Sobolev spaces a standard reference is [Adams and Fournier, 2003]. For very delicate problems one might also consult [Maz'ya, 2011]. More introductory monographs are [Evans and Gariepy, 2015] (for Geometric Measure Theory), [Evans, 2010] (for PDEs, see also the classics: [Gilbarg and Trudinger, 2001], and [Ziemer, 1989]).

7. Normed Vector spaces

There is an order of concepts of spaces

Topological spaces ⊃ metric spaces ⊃ normed vector spaces ⊃ Pre-Hilbert-spaces

Definition 7.1. Let X be a set and $\tau \subset 2^X$ a collection of subsets of X satisfying the following axioms.

(1) $\emptyset \in \tau$ and $X \in \tau$
(2) If $\tau' \subset \tau$ is any arbitrary (finite or infinite, countable or uncountable) union of members of τ, then $A := \bigcup_{B\in\tau'} B$ belongs to τ.
(3) If $(A_k)_{k=1}^N \subset \tau$ then $\bigcap_{k=1}^N A_k \in \tau$ for every finite $N \in \mathbb{N}$.

The elements of τ are called *open sets*. The collection τ is called a *topology* on X.

Definition 7.2. (1) Let X be a set and $d : X \to X \to [0,\infty)$ be a map with the following properties for all $x, y, z \in X$
  (a) $d(x,y) \geq 0$ and equality holds if and only if $x = y$ (non-degeneracy)
  (b) $d(x,y) = d(y,x)$ (symmetry)
  (c) $d(x,y) \leq d(x,z) + d(z,y)$ (triangular inequality)

  Then d is called a *metric* or *distance*, and (X,d) is called a *metric space*.

(2) Let (X,d) be a metric space. For $x \in X$ and $r > 0$ the *open ball* $B(x,r)$ is defined as
$$B(x,r) := \{y \in X : \quad d(x,y) < r\}.$$
  A set $A \subset X$ is called open if for each $x \in A$ there exists $r > 0$ such that $B(x,r) \subset A$.

Exercise 7.3. *Let (X,d) be a metric space and set*
$$\tau := \{A \subset X : \quad A \text{ open}\}$$
Then (X,τ) is a topological space.

τ is sometimes called the associated metric topology*.*

Functional analysis (well: *linear* functional analysis) deals with linear spaces.

Definition 7.4 (linear space)**.** A *linear space*, or *vector space*, $(X, *, +)$ is a set with a scalar multiplication $* : \mathbb{R} \times X \to X$, an vector addition $+ : X \times X$ that satisfy the following properties:

- $u + (v + w) = (u + v) + W$ (associativity of vector addition)
- $u + v = v + u$ (commutativity of vector addition)
- there exists $0 \in X$ (called the origin or zero) such that $v + 0 = v$ for all $v \in V$ (identity element of the group $(V, +)$
- For every $v \in V$ there exists an element $-v \in V$, called the additive inverse of v such that $v + (-v) = 0$ (inverse element)
- for every $\lambda, \mu \in \mathbb{R}$ and $v \in V$ we have $\lambda(\mu v) = (\lambda\mu)v$
- $1v = v$ where $1 \in R$ is the real number 1
- $\lambda(u + v) = \lambda u + \lambda v$ (distributivity)
- $(\lambda + \mu)v = \lambda v + \mu v$ (distributivity).

(we often write λv instead of $\lambda * v$).

A *norm* $\|\cdot\|$ on a linear space X is a map $\|\cdot\| : X \to \mathbb{R}$ with the properties

- $\|v\| \geq 0$ for all $v \in X$ and equality holds if and only if $v = 0$
- $\|\lambda v\| = |\lambda| \, \|v\|$ for all $\lambda \in \mathbb{R}$, $v \in X$
- $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in X$.

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

Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent norms* if there exists $C > 0$ such that

$$C^{-1}\|x\|_2 \leq \|x\|_1 \leq C\|x\|_2 \quad \forall x \in X.$$

Exercise 7.5. *Let $(X, \|\cdot\|)$ be a normed vector space. Show that*

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

is a metric. d is often called the induced metric *for a normed vector space.*

Show that two equivalent norms induce two equivalent metrics.

Definition 7.6. Let $(X, *+)$ be a *linear space*, or *vector space*, $(X, *, +)$. A *scalar product* or *inner product* is a map $\langle \cdot, \cdot \rangle : X \to \mathbb{R}$ with the following properties

(1) $\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in X$
(2) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$
(3) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$ for all $\lambda, \mu \in \mathbb{R}$, $x, y, z \in X$.

A vector space equipped with a scalar product is called a *inner product space* or *pre-Hilbert space*

Exercise 7.7. *Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space and denote*

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

Show that $\|\cdot\|$ is a norm.

$\|\cdot\|$ is then called the induced norm.

Exercise 7.8 (Cauchy-Schwarz). *We have*

$$\langle x, y \rangle \leq \|x\| \, \|y\|$$

equality holds if and only if x and y are linearly dependent*, i.e.*

$$x = \lambda y, \quad \text{or} \quad y = \lambda x$$

for some $\lambda \in \mathbb{R}$.

Hint: Extend the proof of Advanced Calculus.

We can also check if a norm belongs to a scalar product.

Exercise 7.9 (Parallelogram law). *Let $(X, \|\cdot\|)$ be a normed vector space. Then there exists a scalar product $\langle \cdot, \cdot \rangle$ inducing $\|\cdot\|$ if and only if*

$$2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2 \quad \forall x, y \in X.$$

Hint: Set $\langle x, y \rangle := \frac{\|x+y\|^2 - \|x-y\|^2}{4}$.

Definition 7.10. A normed vector space is complete, if it is metrically complete. A normed vector space is called *Banach space*. If space is additionally an inner product space, then we call it *Hilbert space*.

We now recall the definition of the two spaces we want to talk about most of the time in this course $L^p(\Omega)$ and $W^{1,p}(\Omega)$.

Example 7.11. (*L^p-spaces*) Let $\Omega \subset \mathbb{R}^n$ be a μ-measurable set, where for simplicity we assume that μ is a Radon measure (most of the time μ will be the Lebesgue measure). Let $p \in [1, \infty]$. Let $f : \Omega \to \mathbb{R}$ be a μ-measurable function. We say that $f \in L^p(\Omega)$ if and only if

$$\|f\|_{L^p(\Omega,\mu)} := \left(\int_\Omega |f|^p d\mu \right)^{\frac{1}{p}} < \infty.$$

This makes sense only if $p < \infty$, but for $p = \infty$ we set

$$\|f\|_{L^\infty(\Omega,\mu)} := \sup_{x \in \Omega} |f(x)|,$$

where sup is to be understood as the μ-*essential supremum*, i.e.

$$\sup_{x \in \Omega} |f(x)| = \inf\{\Lambda > 0 : \quad \mu\{x \in \Omega : |f(x)| > \Lambda\} = 0\}.$$

$\|\cdot\|_{L^p(\Omega,\mu)}$ is in general not a norm, since $\|f\|_{L^p(\Omega,\mu)} = 0$ implies that $f = 0$ only μ-a.e.

We solve this issue by always assuming that $f = g$ (in the sense of $L^p(\Omega, \mu)$) if $f(x) = g(x)$ for μ-a.e. $x \in \Omega$. (i.e. the set $\mu\{x : f(x) \neq g(x)\} = 0$).

Then $\|\cdot\|_{L^p(\Omega,\mu)}$ is a norm on all $L^p(\Omega, \mu)$-functions (modulo equality). And it is complete, Theorem 3.27.

Also $L^2(\Omega)$ is a Hilbert space with the scalar product

$$\langle f, g\rangle = \int fg d\mu.$$

Example 7.12 (*Sobolev space $W^{1,p}$*)**.** Let $\Omega \subset \mathbb{R}^n$ be any open set. The space $W^{1,p}(\Omega)$ consists of maps $f \in L^p(\Omega)$ such that any first-order distributional derivative $\partial_\alpha f \in L^p(\Omega)$, and equipped with the norm

$$\|f\|_{W^{1,p}(\Omega)} := \|f\|_{L^p(\Omega)} + \|Df\|_{L^p(\Omega)}.$$

Here $\partial_\alpha f \in L^p(\Omega)$ means that (cf. Theorem 4.46)

$$\int_\Omega f\partial_\alpha \varphi = -\int_\Omega g_\alpha \varphi \quad \forall \varphi \in C_c^\infty(\Omega),$$

for some $g_\alpha \in L^p(\Omega)$. We identify $\partial_\alpha f := g_\alpha$, and set $Df = (\partial_1 f, \partial_2 f, \ldots, \partial_n f) \in L^p(\Omega, \mathbb{R}^n)$.

We will later see that this Sobolev space is complete, Proposition 13.8, and indeed it is equivalent to the following *metric closure* of smooth functions under the $W^{1,p}$-norm, (observe that $C^\infty(\Omega)$ and $C^\infty(\overline{\Omega})$ are two different spaces!),

$$W^{1,p}(\Omega) = \overline{\left\{f \in C^\infty(\Omega) : \quad \|f\|_{L^p(\Omega)} + \|Df\|_{L^p(\Omega)} < \infty\right\}}^{\|\cdot\|_{W^{1,p}(\Omega)}}$$

(We have proven this for $\Omega = \mathbb{R}^n$ in Theorem 4.46, we will prove this for any open set in Theorem 13.16).

Let us remark that often (not always) $H^{1,p}$ is used instead of $W^{1,p}$ (H being for Hardy, W being for Weill). And H^1 often (not always) refers to $W^{1,2}$.

$W^{1,2}$ is a Hilbert space, if we use the following norm (which is equivalent to the original norm)

$$\left(\|f\|^2_{L^2(\mathbb{R}^n)} + \|\nabla f\|^2_{L^2(\mathbb{R}^n)}\right)^{\frac{1}{2}},$$

i.e. the following scalar product

$$\langle f, g\rangle_{W^{1,2}} := \int fg + \int \nabla f \cdot \nabla g.$$

Definition 7.13. A vector space X is called *finite* or *finite dimensional* if there exists a finite basis, $X = \operatorname{span}\{b_1, \ldots, b_n\}$ for some $n \in \mathbb{N}$ and some $b_1, \ldots, b_n \in X$, and $(b_i)_{i=1}^n$ *linear independent*, i.e. whenever $(\lambda_i)_{i=1}^n \subset \mathbb{R}$,

$$\sum_{i=1}^n \lambda_i b_i = 0 \quad \Rightarrow \quad \lambda_i = 0 \quad \forall i = 1, \ldots, n.$$

Equivalently, for any $x \in X$ there exists exactly one sequence $(\lambda_i)_{i=1}^n$ such that

$$x = \sum_{i=1}^n \lambda_i b_i.$$

Exercise 7.14. *Let $(X, \|\cdot\|)$ be a finite dimensional normed vector space with basis $b_1, \ldots, b_n$. I.e. for any $x \in X$ there exists exactly one sequence of $(\lambda_i)_{i=1}^n \subset \mathbb{R}$ such that*

$$x = \sum_{i=1}^n \lambda_i b_i.$$

Define a new norm $\|\cdot\|_2$ on X as follows

$$\|x\|_2 := \sqrt{\sum_{i=1}^n |\lambda_i|^2}.$$

Show that $\|\cdot\|_2$ is indeed a norm on X and show that $\|\cdot\|_2$ is equivalent to $\|\cdot\|$. Namely there exist $\Lambda > 0$ such that

$$\Lambda^{-1}\|x\|_2 \le \|x\| \le \Lambda\|x\|_2 \quad \forall x \in X.$$

Hint: To see $\Lambda^{-1}\|x\|_2 \le \|x\|$ you can use a blowup proof and Bolzano Weierstrass. Assume that there is no such Λ, then there must be a sequence $x_k \in X$

$$\|x_k\|_2 > k\|x_k\|.$$

W.l.o.g. $\|x_k\|_2 = 1$. Use Bolzano-Weierstrass to conclude that x_k is actually converging in X and conclude that the limit $\bar{x}$ satisfies $\|\bar{x}\|_2 = 1$ and $\|\bar{x}\| = 0$ simultaneously, which is a contradiction.

Exercise 7.15. *If X is a finite dimensional vector space and $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms. Show that they are equivalent.*

Hint: Exercise 7.14.

Exercise 7.16. *Show that any normed vector space $(X, \|\cdot\|)$ with countably many $(c_i)_{i=1}^\infty \subset X$ such that $\|c_i\| \le 1$ and*

$$\|c_i - c_j\| \ge \frac{1}{100} \quad \forall i \neq j.$$

is infinite dimensional.

Hint: Assume not, then we can assume (7.15) that $X = \mathbb{R}^n$ for some n. Find a contradiction to Bolzano-Weierstrass.

Exercise 7.17. *Let $(X, \|\cdot\|)$ be a normed vector space (identified with its metric space).*

Show that

$$X \text{ is finite dimensional} \quad \Leftrightarrow \quad \text{any closed and bounded set in } X \text{ is compact.}$$

Hint: For one direction you can use the Riesz' Lemma, Lemma 7.18 to apply Exercise 7.16.

Lemma 7.18. *Let $U \subsetneq X$ be a closed subspace of a normed vector space X. For any $\delta \in (0,1)$ there exists $x_\delta \in X$, $\|x_\delta\| = 1$, such that*

$$\|x_\delta - u\| \geq \delta \quad \forall u \in U.$$

Exercise 7.19. *Find proof online for Lemma 7.18*

Exercise 7.20. *In a normed space of infinite dimension no closed ball is compact.*

Exercise 7.21. *Let $\Omega \subset \mathbb{R}^n$ be any nonempty open set and $1 \leq p \leq \infty$. Show that $L^p(\Omega)$ and $W^{1,p}(\Omega)$ are both infinite dimensional.*

Hint: You can construct functions as in Exercise 7.16 using e.g. cutoff functions on small sets.

8. Linear operators, Dual space

We already defined the notion of bounded linear operators $L : X \to Y$ between two normed vector spaces X and Y, see Theorem 5.33.

The linear operator $L : X \to Y$ is *continuous* (often also called *bounded*), if and only if the *operatornorm* $\|L\|_{\mathcal{L}(X,Y)}$ defined in Definition 5.35 is finite,

$$\|L\|_{\mathcal{L}(X,Y)} := \sup_{\|x\|_X \leq 1} \|Lx\|_Y,$$

or equivalently if for some $\Lambda < \infty$

$$\|Lx\|_Y \leq \Lambda \|x\|_X \quad \forall x.$$

($\|L\|_{\mathcal{L}(X,Y)}$ is then the smallest Λ for which this inequality holds).

Exercise 8.1. *Show that for all finite-dimension vector spaces X, Y any linear operator $L : X \to Y$ is continuous.*

Exercise 8.2. *Show that $T : C^1([0,1]) \to C^0([0,1])$ given by $T\varphi := \varphi'$ is continuous, if C^1 is equipped with the C^1-norm, but discontinuous if we equip C^1 with the C^0-norm.*

Exercise 8.3. *Let $\Omega \subset \mathbb{R}^n$ be an open set. Show that the partial derivative ∂_α*

- *As map from $(W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}(\Omega)}) \to (L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ is a linear continuous operator.*
- *As map from $(C^\infty(\overline{\Omega}), \|\cdot\|_{W^{1,p}(\Omega)}) \to L^p(\Omega)$ is a linear continuous operator.*
- *As map from $(C^\infty(\overline{\Omega}), \|\cdot\|_{L^p(\Omega)}) \to L^p(\Omega)$ is not a linear continuous operator.*
- *As map from $(W^{1,p}(\Omega), \|\cdot\|_{L^p(\Omega)}) \to (L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ is not a linear continuous operator.*

Clearly if L and T are two linear operators from X to Y we can define for $\lambda, \mu \in \mathbb{R}$

$$\lambda L + T\mu : X \to Y; \quad (\lambda L + \mu T)(x) = \lambda L(x) + \mu T(x) \in Y.$$

That is the space of bounded linear operators from X to Y is a linaer space, and we denote it by $\mathcal{L}(X,Y)$.

Exercise 8.4. *Show that the operatornorm $\|\cdot\|$ is indeed a norm on this space*

Exercise 8.5. *Let $(X,\|\cdot\|_{X,1})$, $(Y,\|\cdot\|_{Y,1})$ be a normed space and assume that $\|\cdot\|_{X,2}$ is an equivalent norm to $\|\cdot\|_{X,1}$ and $\|\cdot\|_{Y,2}$ is an equivalent norm to $\|\cdot\|_{Y,1}$.*

(1) *Show that any linear bounded function from $X \to Y$ w.r.t the first norms is also linear bounded function with respect to the second norm.*
(2) *Show that the operator norms with respect to the first norms is equivalent to the operatornorm w.r.t second norms.*

That is $\mathcal{L}(X,Y)$ is independent of the specific (equivalent!) choice of norms.

Exercise 8.6. *Let X,Y,Z be normed spaces and $T \in \mathcal{L}(X,Y)$, $S \in \mathcal{L}(Y,Z)$. Define*

$$S \circ T(X) := S(T(x)).$$

Show that $S \circ T \in \mathcal{L}(X,Z)$ and

$$\|S \circ T\|_{\mathcal{L}(X,Z)} \leq \|S\|_{\mathcal{L}(Y,Z)} \|T\|_{\mathcal{L}(X,Y)}.$$

Definition 8.7. The space of bounded linear operators from X to Y equipped with the operatornorm is denoted by $\mathcal{L}(X,Y)$. The operator norm is usually denoted by $\|\cdot\|_{\mathcal{L}(X,Y)}$.

If $Y = \mathbb{R}$ then we write $X^* := \mathcal{L}(X,\mathbb{R})$, and call X^* the *dual space* to X. The operator norm is $\|\cdot\|_{X^*}$.

Exercise 8.8. *Assume that a normed space $(X,\|\cdot\|_X)$ can be written as $X = X_1 \times X_2$ where $(X_i,\|\cdot\|_{X_i})$ are two normed spaces. Assume furthermore that the norm $\|\cdot\|_X$ is equivalent to*

$$X = X_1 \times X_2 \ni (x_1,x_2) \mapsto \|x_1\|_{X_1} + \|x_2\|_{X_2}$$

Show that

$$X^* = X_1^* \oplus X_2^*,$$

in the following sense

(1) *every $x^* \in X^*$ can be written as*
$$x^*[(x_1,x_2)] = x_1^*(x_1) + x_2^*(x_2).$$
where $x_i^* \in X_i^*$ for $i = 1,2$.
(2) *every pair $x_i^* \in X_i^*$, $i = 1,2$, induce an element $x^* \in X^*$ via*
$$x^*[(x_1,x_2)] = x_1^*(x_1) + x_2^*(x_2).$$
(3) *Show that x_1^* and x_2^* are uniquely determined, i.e. if*
$$x_1^*(x_1) + x_2^*(x_2) = \tilde{x}_1^*(x_1) + \tilde{x}_2^*(x_2) \forall (x_1,x_2) \in X$$
then $x_i^* = \tilde{x}_i^*$ for $i = 1,2$.
(4) *The norms $\|x^*\|_{X^*}$ and $\|x_1^*\|_{X_1^*} + \|x_2^*\|_{X_2^*}$ are equivalent.*

Theorem 8.9. *If X is a normed vector space and Y is complete, then $\mathcal{L}(X, Y)$ is complete. In particular, X^* is always complete.*

Proof. Let T_n be a Cauchy-Sequence in $L(X, Y)$. Then for each fixed x, $(T_n x)_n \subset Y$ is a Cauchy sequence, and thus we can define

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

It is easy to check that T is now linear and continuous. (exercise!) □

Example 8.10. We have discussed in Theorem 5.39 (Riesz Representation Theorem) that for $1 \le p < \infty$ any element in $f^* \in (L^p(\mathbb{R}^n))^*$ can be represented by some $f \in L^{p'}(\mathbb{R}^n)$ (recall $p' = \frac{p}{p-1}$ is the Hölder dual!), via

$$f^*[\varphi] = \int_{\mathbb{R}^n} f\varphi.$$

A possible element in $(W^{1,p}(\mathbb{R}^n))^*$ is

$$f^*[\varphi] := \int_{\mathbb{R}} f_1\varphi + \int_{\mathbb{R}} f_2 \cdot D\varphi,$$

wher $f_1 \in L^p(\mathbb{R}^n)$ and $f_2 \in L^p(\mathbb{R}^n, \mathbb{R}^n)$. We will need Hahn-Banach theorem, Theorem 10.2 below, to show that this is the generic form of a dual element, Corollary 10.12.

8.1. **Compact operators.** Let us also define *compact operator*.

Definition 8.11.
- Let (X, d) be a metric space. A set $A \subset X$ is called *precompact*, if any sequence $(a_k)_{k\in\mathbb{N}} \subset A$ has a convergent subsequence *in X*.
- Let (X, d_X) and (Y, d_Y) be two metric spaces. A compact operator T is called a *compact operator* if it maps bounded sets $A \subset X$ to precompact sets $T(A) \subset Y$[27].

Exercise 8.12. *Show that a set $A \subset X$ is pre-compact if and only if its closure $\overline{A}$ is compact.*

Exercise 8.13.
- *Show that in general, a compact operator $L : (X, d_x) \to (Y, d_Y)$ may not be continuous.*
- *Show that if (X, d_x) and (Y, d_Y) are normed vector spaces (with d_x deriving from the X-norm, and d_y deriving from the Y-norm) then any compact linear operator $T : X \to Y$ is actually continuous.*

Compact operators are somewhat lower order, cf. Theorem 13.51. As we shall later see, the identity map $I : (W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}(\Omega)}) \to (L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ is compact if Ω is a smoothly bounded set, Theorem 13.35, which is called the *Rellich-Kondrachov Theorem*. That is whenever $(f_k)_{k\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega)$, i.e. $\sup_k \|f_k\|_{W^{1,p}} < \infty$ there exists a subsequence $(f_{k_i})_{i\in\mathbb{N}}$ such that f_{k_i} converges in $L^p(\Omega)$.

[27]Careful: for general metric spaces this definition may not be uniform in the literature, some people might assume that compact operators are continuous – we don't care about this for linear operators, Exercise 8.13

We also have a version of this for continuous maps, based on Arzela-Ascoli (and indeed, the Rellich-Kondrachov Theorem is a consequence of Arzela-Ascoli as well).

Example 8.14. Let $\overline{\Omega} \subset \mathbb{R}^n$ be a compact set. Denote by $Y := C^0(\overline{\Omega})$ the set of continuous functions on $\overline{\Omega}$ equipped with the supremums-norm

$$\|f\|_Y := \|f\|_{L^\infty(\Omega)}.$$

Denote by $X := C^{0,1}(\overline{\Omega})$ the set of uniformly Lipschitz continuous functions on $\overline{\Omega}$ equipped with the norm

$$\|f\|_X := \|f\|_{L^\infty(\Omega)} + \sup_{x \neq y \in \overline{\Omega}} \frac{|f(x) - f(y)|}{|x-y|}.$$

Denote by $I : X \to Y$ the identity map, $If = f$. It is obviously linear, and it is bounded, since

$$\|If\|_Y = \|f\|_{L^\infty(\Omega)} \leq \|f\|_{L^\infty(\Omega)} + \sup_{x \neq y \in \overline{\Omega}} \frac{|f(x) - f(y)|}{|x-y|} = \|f\|_X$$

Indeed, it $I : X \to Y$ is a compact operator: take $(f_k)_{k\in\mathbb{N}}$ a bounded sequence in X, i.e.

$$\sup_k \|f_k\|_X < \infty.$$

From the definition of $\|X\|$ we observe that $(f_k)_{k\in\mathbb{N}}$ is then uniformly bounded and equicontinuous. By Arzela-Ascoli we know that then there exists a uniformly converging subsequence, i.e. $(f_{k_i})_{i\in\mathbb{N}}$ and $f \in C^0(\overline{\Omega})$ such that

$$\|f_{k_i} - f\|_{L^\infty} \xrightarrow{i\to\infty} 0.$$

But this is the same as to say that $(If_{k_i})_{i\in\mathbb{N}}$ is convergent in X. That is, $I : X \to Y$ maps bounded sets in X to pre-compact sets in Y.

9. Subspaces and Embeddings

Definition 9.1. Let X and Y be two normed spaces.

- A linear vector space X is *embedded* into a linear vector space Y, if there exists a map $T : X \to Y$ which is bounded linear and injective. We then say that X is embedded into Y (via T), in symbols.

$$X \overset{T}{\hookrightarrow} Y$$

- We say that $X \overset{T}{\hookrightarrow} Y$ is an *isometric embedding* if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$.
- We say the embedding is $X \overset{T}{\hookrightarrow} Y$ *compact*, or X is *compactly embedded* in Y if the embedding operator T is compact.
- If $T : X \to Y$ is the identity, and X and Y carry the same norm, then we say X is a (normed) *subspace* of Y.

Example 9.2. $C^{0,1}(\overline{\Omega})$ is a subspace of $C^0(\overline{\Omega})$, but in view of Example 8.14 it is also compactly embedded.

What is the issue with these statements? *The norms are not specified!*

We should have written that $(C^{0,1}(\overline{\Omega}), \|\cdot\|_{L^\infty(\Omega)})$ is a subspace of $\left(C^0(\overline{\Omega}), \|\cdot\|_{L^\infty(\Omega)}\right)$ (nobody does that, though).

And, if we set

$$\|f\|_{C^{0,1}} := \|f\|_{L^\infty(\Omega)} + \sup_{x\neq y\in\overline{\Omega}} \frac{|f(x)-f(y)|}{|x-y|},$$

then $(C^{0,1}(\overline{\Omega}), \|\cdot\|_{C^{0,1}})$ is *not* a subspace of $\left(C^0(\overline{\Omega}), \|\cdot\|_{L^\infty(\Omega)}\right)$, but it is compactly embedded (via the identity map).

Exercise 9.3. *For any $p \in [1,\infty)$, $\Omega \subset \mathbb{R}^n$ open:*

- *$(W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}(\Omega)})$ is embedded into $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ (under the identity map, but this is not a isometric embedding).*
- *$(W^{1,p}(\Omega), \|\cdot\|_{L^p(\Omega)})$ is a subspace of $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$ but show that $(W^{1,p}(\Omega), \|\cdot\|_{L^p(\Omega)})$ is not a closed subspace (Hint: Exercise 9.4)*

Exercise 9.4. *Let Ω be any open set in $\mathbb{R}^n$.*

Show that $(W^{1,p}(\Omega), \|\cdot\|_{L^p(\Omega)})$ is a dense subspace of $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$.

Hint: Use that we have shown that $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$, Theorem 3.32.

As mentioned before, we will later show the Rellich-Kondrachov theorem, Theorem 13.35, that shows that $(W^{1,p}(\Omega), \|\cdot\|_{W^{1,p}})$ is compactly embedded in $L^p(\Omega)$ for nicely bounded sets Ω.

Example 9.5. For $\Omega \subset \mathbb{R}^n$, we can also embed $(W^{1,p}(\Omega), \|\cdot\|_{W1,p(\Omega)})$ *isometrically* into $(L^p(\Omega))^{n+1} = L^p(\Omega) \times \ldots L^p(\Omega)$.

Indeed, let u (for simplicity, easy to change for the original definition (exercise: whats the difference?)) take the $W^{1,p}$-norm to be

$$\|f\|_{W^{1,p}(\Omega)} := \|f\|_{L^p(\Omega)} + \sum_{\alpha=1}^n \|\partial_\alpha f\|_{L^p(\Omega)}.$$

Then let

$$Tf := (f, \partial_1 f, \partial_2 f, \ldots, \partial_n f).$$

Clearly $T : (W^{1,p}(\Omega), \|\cdot\|_{W1,p(\Omega)}) \to (L^p(\Omega))^{n+1}$ is injective. Moreover

$$\|Tf\|_{L^p(\Omega)\times \ldots L^p(\Omega)} = \|f\|_{W^{1,p}(\Omega)}.$$

So T is an isometric embedding.

Exercise 9.6. *Let $\overline{\Omega} \subset \mathbb{R}^n$ be a compact set. Show that $C^{0,\alpha}(\overline{\Omega})$ is compactly embedded in $C^{0,\beta}(\overline{\Omega})$, whenever $0 \leq \alpha \leq \beta \leq 1$, in the following sense:*

Denote by $Y := C^{0,\gamma}(\overline{\Omega})$ the set of Hölder continuous functions, i.e. f such that

$$\sup_{x \neq y \in \overline{\Omega}} \frac{|f(x) - f(y)|}{|x-y|^\gamma} < \infty \quad \text{if } \gamma > 0$$

and f simply continuous if $\gamma = 0$.

Denote by

$$\|f\|_{C^{0,\gamma}(\overline{\Omega})} := \|f\|_{L^\infty(\Omega)} + \sup_{x \neq y \in \overline{\Omega}} \frac{|f(x) - f(y)|}{|x-y|^\gamma}.$$

Show that for $0 \leq \alpha \leq \beta \leq 1$ the identity map

$$I: \quad \left(C^{0,\beta}(\overline{\Omega}), \|\cdot\|_{C^{0,\beta}(\overline{\Omega})}\right) \to \left(C^{0,\alpha}(\overline{\Omega}), \|\cdot\|_{C^{0,\alpha}(\overline{\Omega})}\right)$$

is a compact operator.

Exercise 9.7. *Let $U \subset X$ be a dense subspace, and $T \in \mathcal{L}(U, Y)$ for some Banach space Y. Then there exists a unique extension $S \in L(X,Y)$ with $S\Big|_U = T$. Moreover*

$$\|S\|_{\mathcal{L}(X,Y)} = \|T\|_{\mathcal{L}(U,Y)}$$

Exercise 9.8. *For $\Omega \subset \mathbb{R}^n$ be a bounded set (with smooth boundary, say a ball – not so important for our point here).*

Let $U := C^1(\overline{\Omega})$ be the set of continuously differentiable functions and $Y := L^p(\Omega)$ (with respect to the Lebesgue measure). Fix $\alpha \in \{1, \ldots, n\}$. Define

$$T : U \to Y$$

be defined as (classical derivative!)

$$Tf := \partial_\alpha f.$$

(Let us assume) we know that U is dense in $L^p(\Omega)$ w.r.t. to the L^p-norm, and U is dense in $W^{1,p}(\Omega)$ w.r.t $W^{1,p}$-norm.

Considering Exercise 9.7,

- *what is the extended operator S with respect to $X = W^{1,p}(\Omega)$ (and the $W^{1,p}$-norm),*
- *what is the extended operator S with respect to $X = L^p(\Omega)$ (and the L^p-norm)*

Exercise 9.9. *Let Y be a normed space and $X \subset Y$ be a (linear) subspace. If Y is a Banach space, then X is a Banach space if and only if X is (metrically) closed.*

10. Hahn-Banach Theorem

Example 10.1. Let X be a subspace of Y then Y^* is a subspace of X^* in the following sense. If $y^* \in Y^*$ then clearly $y^*\Big|_X$ is a linear bounded operator on X, and in that sense $y^* \in X^*$.

If X is embedded into Y via the map $T : X \to Y$, then Y^* is embedded into X^* under the operator T^* defined as follows:

$$T^*(y^*)(x) := y^*(T(x)).$$

On the other hand, in the situation of Example 10.1, the following Hahn-Banach theorem tells us, than any $x^* \in X^*$ can be (non-uniquely!) extended to an element of Y^*.

Theorem 10.2 (Hahn-Banach theorem)**.** *Let X be a vector space over $\mathbb{R}$, $U \subset X$ a linear subspace, and let $p : X \to \mathbb{R}$ be sublinear, that is*

(1) $p(\lambda x) = \lambda p(x)$ for all $x \in X$ and $\lambda \geq 0$, $\lambda \in \mathbb{R}$
(2) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$

Assume $T : U \to \mathbb{R}$ be linear and $T(x) \leq p(x)$ for all $x \in U$.

Then there exists a linear extension $T^e : X \to \mathbb{R}$, i.e. a linear map with

(1) $T^e(x) = T(x)$ for all $x \in U$
(2) $T^e(x) \leq p(x)$ for all $x \in X$.

Exercise 10.3. *Show that Theorem 10.2 can be applied to bounded operators. Namely,*

- *Show that $p(x) = \|x\|$ is sublinear.*
- *Assume $\Lambda \in \mathbb{R}$, X is a normed vector space, and T is linear operator $T : X \to \mathbb{R}$ with $Tx \leq \Lambda \|x\|$ for all $x \in X$. Show that $\|T\|_{X^*} \leq \Lambda$, i.e. T is bounded.*

We have already proven the finite-dimensional version in Advanced Calculus (see my Lecture Notes of Adv.Calc, Lemma 28.7.), namely we have

Proposition 10.4. *Let X be a vector space over $\mathbb{R}$, $U \subset X$ a linear subspace and $v \in X \backslash U$. Let $p : X \to \mathbb{R}$ be sublinear, that is*

(1) $p(\lambda x) = \lambda p(x)$ for all $x \in X$ and $\lambda \geq 0$, $\lambda \in \mathbb{R}$
(2) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$

Assume $T : U \to \mathbb{R}$ be linear and $T(x) \leq p(x)$ for all $x \in U$.

Set

$$W := \operatorname{span}\{U, v\} = \{w \in X : w = u + \lambda v \quad \text{for some } \lambda \in \mathbb{R},\ u \in U\}.$$

Then there exists a linear extension $T^e : W \to \mathbb{R}$*, i.e. a linear map with*

(1) $T^e(x) = T(x)$ *for all* $x \in U$
(2) $T^e(x) \leq p(x)$ *for all* $x \in W$.

Proposition 10.4 is enough to prove Theorem 10.2 if X is finite dimensional. But for general linear spaces (which might not even have a countable basis) we need *Transfinite Induction*, i.e. the axiom of choice, here in the form (and equivalent to) the Lemma of Zorn.

Definition 10.5 (Partial Order)**.** Let P be a set.

(1) A map $\leq : A \subset X \times X \to \{true, false\}$ is called a *partial order* (and then X is *partially ordered*) if the following holds. Here and henceforth we follow the notion: we will write $x \leq y$ if $(x, y) \in A$ and $\leq (x, y) = true$, and $x \geq y$ if $x \leq y$. However, observe that there might be $x, y \in X$ such that neither $x \leq y$ nor $y \leq x$.
  - (reflexive): for all $x \in X$ we have $x \leq x$
  - (antisymmetric): If $x \leq y$ and $y \leq x$ then $y = x$.
  - (transitive): If $x \leq y$ and $y \leq z$ then $x \leq z$.

(2) If moreover for every $x, y \in X$ we have either $x \leq y$ or $y \leq x$ or both, then we say that X is *totally ordered*.
(3) If $(P, \leq)$ is a partially ordered set and $S \subset P$, then S is a *chain* if $(S, \leq)$ is totally ordered (where of course $\leq$ is restricted to $S \times S$)
(4) If $(P, \leq)$ is a partially ordered set and $S \subset P$ then $u \in P$ is an *upper bound* of S if $s \leq u$ for all $s \in S$.
(5) If $(P, \leq)$ is a partially ordered set then $m \in P$ is *maximal* if there is no element larger, i.e. for any $s \in P$ with $s \geq m$ we have $s = m$. (but there may be elements which cannot be compared!)

Exercise 10.6. *Let* Y *be any set and*

$$X := 2^Y = \{A \subset Y\}$$

the set of subsets of Y *(i.e. the* power set *of* Y*). Show that the set-inclusion* $\subset$ *is a partial order, but* $(X, \subset)$ *is not necessarily totally ordered.*

Lemma 10.7 (Zorn's Lemma)**.** *Let* $(P, \leq)$ *be a nonempty partially ordered set. Assume that every chain* A *in* P *has an upper bound (in* P*, not necessarily in* A*). Then the set* P *contains at least one maximal element.*

Proof of Theorem 10.2. If X is finite dimensional, Theorem 10.2 follows by induction from Proposition 10.4.

The general case is more abstract, using Zorn's lemma. Let A be the set of collections of extensions of T, i.e.

$$A := \left\{(\tilde{T}, V) : \quad \tilde{T} : V \to \mathbb{R} \text{ is linear}, \tilde{T}\Big|_U = T, \quad \tilde{T}(x) \leq p(x)\ \forall x \in V\right\}.$$

(Here V is always a linear subspace).

We can equip A with the partial order $\leq$, namely

$$(T_1, V_1) \leq (T_2, V_2) \quad :\Leftrightarrow \quad V_1 \subset V_2 : \quad T_2\Big|_{V_1} = T_1.$$

Then A is nonempty (since $(T, U) \in A$).

We want to apply Zorn's lemma, Lemma 10.7, to A. So let $B \subset A$ be a chain, i.e. a totally ordered subset of A – that is for any $(T_1, V_1), (T_2, V_2) \in B$ we have either $(T_1, V_1) \leq (T_2, V_2)$ or $(T_1, V_1) \geq (T_2, V_2)$.

We need to find an upper bound for B in A. For this set

$$\bar{V} := \bigcup_{(\tilde{T}, V) \in B} V.$$

It is easy to check that $\bar{V}$ is a linear subspace. Also we can define $\bar{T} : \bar{V} \to \mathbb{R}$. Let $x \in \bar{V}$, then there must be at least one $(\tilde{T}, V) \in B$ such that $x \in V$. We then set $\bar{T}(x) = \tilde{T}(x)$. If there is any other $(\tilde{T}_2, V_2) \in B$ such that $x \in V_2$ we have (by the assumption of total order) that $\tilde{T}(x) = \tilde{T}_2(x)$ (since one is the extension of the other).

So we have found $\bar{T} : \bar{V} \to \mathbb{R}$. Now let $x, y \in \bar{V}$, $\lambda, \mu \in \mathbb{R}$. Then there exists $(V_1, T_1) \in B$ such that $x \in V_1$, and $(V_2, T_2) \in B$ such that $y \in V_2$. By the assumption of total order of B, we have either $V_1 \subset V_2$ or $V_1 \supset V_2$. Lets assume that $V_1 \subset V_2$ then $\lambda x + \mu y \in V_2$, and we have

$$\bar{T}(\lambda x + \mu y) = T_2(\lambda x + \mu y) = \lambda T_2 x + \mu T_2 y = \lambda \bar{T} x + \mu \bar{T} y.$$

Thus $\bar{T}$ is linear.

If $x \in \bar{V}$ then $\bar{T}x = \tilde{T}x$ for some $(\tilde{T}, V) \in B$, and thus

$$\bar{T}x = \tilde{T}x \leq p(x).$$

Thus, $(\bar{V}, \bar{T}) \in A$. Next we claim that $(\bar{V}, \bar{T})$ is an upper bound for B. And indeed if $(\tilde{T}, V) \in B$, then $V \subset \bar{V}$ and $\bar{T}x = \tilde{T}x$.

So we have established the conditions for the Lemma of Zorn, Lemma 10.7, and conclude that there must be some maximal element $(T^e, W) \in A$, i.e. whenever $(T^e, W) \leq (\tilde{T}, V)$ then $V = W$ and $\tilde{T} = T^e$.

In particular $T^e : W \to \mathbb{R}$ is linear and $T^e\Big|_U = T$ and $T^e(x) \leq p(x)$. All that remains to show is that $W = X$. By contradiction, assume this is not the case. Then there exists $v \in X \backslash W$ to which we can apply Proposition 10.4. We then find an $\tilde{T}^e : \operatorname{span}(v, W) \to \mathbb{R}$ which satisfies all the properties of an element in A. But since T^e is a maximal element we have $W \supset \operatorname{span}(v, W)$ – a contradiction since $v \notin W$. We conclude that $W = X$ and T^e is the extension we wanted. $\square$

Remark 10.8. Observe that the extenstion T^e is in no way claimed to be unique. If in Theorem 10.2 U is dense, we know there is a unique extension to X – but for non-dense U there is no reason that would be true.

Hahn-Banach has several re-formulations and consequences

Corollary 10.9. *Let X be a normed space and $Y \subset X$ be a subspace and $u^* \in Y^*$. Then there exists $x^* \in X^*$ such that*

$$x^*[x] = u^*[x] \quad \forall x \in Y$$

and

$$\|x^*\|_{X^*} = \|u^*\|_{Y^*}.$$

(conpare this with Example 10.1)

Exercise 10.10. *Prove Corollary 10.9*

Corollary 10.11. *Let X be a normed vector space and $x_0 \in X \setminus \{0\}$. Then there exists $x^* \in X^*$ with $\|x^*\|_{X^*} = 1$ and $x^*[x_0] = \|x_0\|_X$.*

Proof. For $\lambda \in \mathbb{R}$ set

$$Y := \operatorname{span}\{x_0\}$$

and set

$$y^*[\lambda x_0] := \lambda \|x_0\|_X.$$

Then $y^* \in Y^*$, and $\|y^*\|_{Y^*} = 1$.

Now take the Hahn-Banach-extension $x^* : X \to \mathbb{R}$ with $p(v) = \|v\|_X$, Theorem 10.2. Then we have

$$|x^*[v]| \le p(v) = \|v\| \quad \forall v \in X.$$

Thus $\|x^*\|_{X^*} \le 1$. Plugging in $v := x_0$ we find that

$$|x^*[x_0]| = \|x_0\|,$$

so we have $\|x^*\|_{X^*} \ge 1$. We can conclude. □

Corollary 10.12. *Let $1 \le p < \infty$ and let $T \in (W^{1,p}(\mathbb{R}^n))^*$. Then there exists $g \in L^{p'}(\mathbb{R}^n)$ and $G \in L^{p'}(\mathbb{R}^n, \mathbb{R}^n)$ such that*

$$T[f] \mapsto \int gf + \int G \cdot Df.$$

Moroever for some $C > 0$ depending only on the dimension n

$$C^{-1} \|T\|_{(W^{1,p}(\mathbb{R}^n))^*} \le \|g\|_{L^{p'}(\mathbb{R}^n)} + \|G\|_{L^{p'}(\mathbb{R}^n,\mathbb{R}^n)} \le C \|T\|_{(W^{1,p}(\mathbb{R}^n))^*}.$$

Proof. Let

$$X := \underbrace{L^p(\mathbb{R}^n) \times \ldots \times L^p(\mathbb{R}^n)}_{n+1 \text{ times}}$$

with the norm

$$\|(f_0, \ldots, f_n)\|_X := \sum_{i=0}^{n+1} \|f_i\|_{L^p(\mathbb{R}^n)}.$$

Let $Y \subset X$ be given by

$$Y = \{(f_0, f_1 \ldots, f_n) \in X : \quad f_i = \partial_i f_0, \text{ i=1}, \ldots, n\}$$

where the derivative is taken in the sense of distribution. Clearly Y is linear subspace of X[28]

Let $T \in (W^{1,p}(\mathbb{R}^n))^*$, then we can consider T as a bounded linear functional on Y, i.e. $T \in Y^*$. By Hahn-Banach theorem, Corollary 10.9, we can extend T to $T^e \in X^*$. In view of Exercise 8.8, we have that

$$X^* = (L^p(\mathbb{R}^n))^* \oplus (L^p(\mathbb{R}^n))^* \oplus \ldots \oplus (L^p(\mathbb{R}^n))^*.$$

Applying componentwise the Riesz-representation theorem, Theorem 5.39, we find $g_0, \ldots, g_n \in L^{p'}(\mathbb{R}^n)$ such that

$$T^e(f_0, f_1, \ldots, f_n) = \sum_{i=0}^{n} \int_{\mathbb{R}^n} g_i f_i.$$

If we set $G := (g_1, \ldots, g_n)^t$ then in particular for $f \in W^{1,p}(\mathbb{R}^n)$ we have

$$T(f) = T^e(f, \partial_1 f, \ldots, \partial_n f) = \int_{\mathbb{R}^n} g_0 f + \int_{\mathbb{R}^n} G \cdot Df.$$

□

Remark 10.13. Actually, one can sharpen Corollary 10.12, and show that

$$T[f] \mapsto \int gf + \int Dg_2 \cdot Df.$$

for some $g_2 \in \dot{W}^{1,p}(\mathbb{R}^n)$. One can prove this by so-called *Hodge decomposition* (also sometimes referred to as *Helmholtz decomposition*), which says that we can split $G = Dg_2 + \tilde{G}$ where $\tilde{G}$ is divergence free (and thus $\int \tilde{G} \cdot Df$ vanishes. The construction of G be done variationally, by minimizing $g_2 \mapsto \|G - Dg_2\|_{L^2(\mathbb{R}^n)}$, but the L^p-estimates (if $p \neq 2$) need Calderon-Zygmund theory (i.e. Harmonic Analysis).

[28] Actually Y is also closed. Observe that

$$\mathcal{I} : f \mapsto (f, \partial_1 f, \ldots, \partial_n f)$$

is then an linear map from $W^{1,p}(\mathbb{R}^n)$ to Y. It is clearly injective and onto. Moreover we have

$$\|\mathcal{I} f\|_X = \|f\|_{W^{1,p}(\mathbb{R}^n)}.$$

Since $W^{1,p}(\mathbb{R}^n)$ is complete, Y is a closed linear subspace of X.

Corollary 10.14. *Let X be a normed vector space. Then*

$$\|x\|_X = \max_{x^* \in X^*, \|x^*\|_{X^*}=1} |x^*[x]|$$

(Observe that the maximum is obtained)

Proof. Fix $x \in X$. Clearly we have

$$\sup_{x^* \in X^*, \|x^*\|_{X^*}=1} |x^*[x]| \le \|x\|_X.$$

Now take $\bar{x}^*$ from Corollary 10.11. Then

$$\sup_{x^* \in X^*, \|x^*\|_{X^*}=1} |x^*[x]| \ge |\bar{x}^*[x]| = \|x\|_X.$$

Thus we have

$$\sup_{x^* \in X^*, \|x^*\|_{X^*}=1} |x^*[x]| = |\bar{x}^*[x]| = \|x\|_X$$

In particular the supremum is attained. □

Exercise 10.15. *Let X be a normed vector space and $x \in X$. Show that: $x^*[x] = 0$ for all $x^* \in X^*$ implies $x = 0$.*

Slighly more generally than Corollary 10.14

Exercise 10.16. *Prove the following. Let X be a normed vector space and let $U \subset X^*$ be a dense set. Then*

$$\|x\|_X = \sup_{u^* \in U, \|u^*\|_{X^*} \le 1} |u^*[x]|$$

A specific application of Exercise 10.16 and the Riesz-Representation theorem is the following *duality* argument.

Proposition 10.17. *Let $p \in [1, \infty]$ then*

$$\|f\|_{L^p(\mathbb{R}^n)} = \sup_{g \in C_c^\infty(\mathbb{R}^n), \|g\|_{L^{p'}} \le 1} \int fg$$

Actually neither Hahn-Banach nor Riesz representation theorem is needed for Proposition 10.17, one can argue as in Equation (5.10).

We can also use now functionals to seperate subspaces, which will be very important for the reflexivity of $W^{1,p}$ later, Theorem 11.9 and Corollary 11.10.

Corollary 10.18. *Let X be a normed vector space and $U \subset X$ a subspace which is additionally closed. Assume $x_0 \in X \setminus U$. Then there exists $x^* \in X^*$ such that $x^*[x] = 0$ for all $x \in U$, but $x^*[x_0] = 1$.*

Proof. Denote

$$V := \operatorname{span}(U, x_0).$$

Since U is a subspace and $x_0 \notin U$, for any $v \in V$ there exists exactly one $u \in U$ and one $\lambda \in \mathbb{R}$ such that

$$v = u + \lambda x_0.$$

Indeed, assume we have $\tilde{u} + \tilde{\lambda} x_0 = u + \lambda x_0$ then we have

$$\tilde{u} - u = \left(\lambda - \tilde{\lambda}\right) x_0$$

If $\tilde{\lambda} \neq \lambda$ we obtain that

$$U \ni \frac{\tilde{u} - u}{\lambda - \tilde{\lambda}} = x_0 \notin U$$

So $\tilde{\lambda} = \lambda$ and thus $\tilde{u} = u$ and we have shown uniqueness.

Now we define the operator on V. For $v = u + \lambda x_0$ set

$$Tv := \lambda.$$

Clearly T is linear. It remains to show that $T \in V^*$.

$$|Tv| \leq |\lambda|. \tag{10.1}$$

So, what we need to show is

$$|\lambda| \leq \|u + \lambda x_0\|. \tag{10.2}$$

Observe that since U is a linear space,

$$\|u + \lambda x_0\|_X = |\lambda| \| \underbrace{-\frac{1}{\lambda} u}_{\in U} - x_0\|_X \geq |\lambda| \inf_{\tilde{u} \in U} \|\tilde{u} - x_0\|.$$

So all we need to show is

$$\inf_{\tilde{u} \in U} \|\tilde{u} - x_0\| > 0. \tag{10.3}$$

This is the place where the closedness of U comes into play. Assume

$$\inf_{\tilde{u} \in U} \|\tilde{u} - x_0\| = 0.$$

Then there exists $\tilde{u}_k \in U$ such that $\|\tilde{u}_k - x_0\| \xrightarrow{k \to \infty} 0$. That is $x_0 \in \overline{U}$. Since U is closed we would have $x_0 \in U$, a contradiction.

Thus (10.3) is established, which implies (10.2), which in view of (10.1) implies

$$|Tv| \leq C\,|v|$$

where $C = (\inf_{\tilde{u} \in U} \|\tilde{u} - x_0\|)^{-1}$. Thus $T \in V^*$ and we can use the Hahn-Banach extension to conclude. $\square$

Recall the definition of separable spaces

Definition 10.19. A normed space X is *separable* iff there exists a countable set $U = \{x_1, x_2, \ldots\}$ with $\overline{U} = X$.

Corollary 10.20. *Let X be a normed vector spcae. If X^* is separable, then X is separable*

Proof. Let $x_1^*, x_2^*, \ldots \subset X^*$ be a dense sequence.

For each x_i^* there must be some $x_i \in X$, $\|x_i\|_X = 1$ such that

$$x_i^*[x_i] \geq \frac{1}{2}\|x_i^*\|_{X^*}.$$

$$Y_{\mathbb{Q}} := \left\{ x \in X : x = \sum_{i=1}^n \lambda_i x_i, \quad \text{for some } n \in \mathbb{N} \text{ and } (\lambda_i)_{i=1}^n \in \mathbb{Q} \right\}$$

and

$$Y_{\mathbb{R}} := \left\{ x \in X : x = \sum_{i=1}^n \lambda_i x_i, \quad \text{for some } n \in \mathbb{N} \text{ and } (\lambda_i)_{i=1}^n \in \mathbb{R} \right\}$$

Clearly $Y_{\mathbb{Q}}$ is countable, and it is dense in $Y_{\mathbb{R}}$. So all we need to show is that the closure $\overline{Y_{\mathbb{R}}} = X$. Assume this is not the case then there exists $x_0 \in X \setminus \overline{Y_{\mathbb{R}}}$. $\overline{Y_{\mathbb{R}}}$ is a closed linear subspace (here we use $Y_{\mathbb{R}}$, otherwise we could have worked with $(x_i)_i$). So by Hahn-Banach, Corollary 10.18, there exists $x_0^* \in X^*$ such that $x_0^*[x] = 0$ for all $x \in \overline{Y}_{\mathbb{R}}$ but $x_0^*[x_0] = 1$.

We then have for each $n \in \mathbb{N}$

$$\|x_0^* - x_n^*\|_{X^*} \overset{\|x_n\|_X = 1}{\geq} |(x_0^* - x_n^*)(x_n)| \overset{x_n \in Y_R}{=} |x_n^*(x_n)| \geq \frac{1}{2}\|x_n^*\|_{X^*},$$

by the choice of x_n. By reverse triangular inequality this implies

$$\|x_0^* - x_n^*\|_{X^*} \geq \frac{1}{2}\|x_n^*\|_{X^*} \geq \frac{1}{2}\|x_0^*\|_{X^*} - \frac{1}{2}\|x_0^* - x_n^*\|_{X^*},$$

that is

$$3\|x_0^* - x_n^*\|_{X^*} \geq \|x_0^*\|_{X^*}.$$

This holds for any $n \in \mathbb{N}$. Thus

$$3 \inf_{n \in \mathbb{N}} \|x_0^* - x_n^*\|_{X^*} \geq \|x_0^*\|_{X^*}.$$

By density assumption the left-hand side is zero, so we have $x_0^* = 0$, a contradiction to $x_0^*(x_0) = 1$. $\square$

10.1. **Separation theorems.** For any convex set C and any point x_0 outside the convex set C there exists a line that separates C and x_0. This is true in any dimension (straight lines are represented by $x^*(x) = c$, one side of a straight line is $\leq c$, the other one $\geq c$. In infinite dimensions this is a consequence of Hahn-Banach.

Recall that $C \subset X$ is *convex* if and only if for any $x, y \in C$ we have $\lambda x + (1 - \lambda)y \in C$ for all $\lambda \in [0, 1]$.

There are several versions of the *separation theorems* which are of fundamental importance in convex optimization.

Theorem 10.21. *Let X be a normed space and $C \subset X$ nonempty, open, and convex and let $x_0 \in X \setminus C$. Then there exists $x^* \in X^*$ such that*

$$x^*[x] < x^*[x_0] \quad \forall x \in C.$$

Proof. We may assume that $0 \in C$. Otherwise let $\tilde{C} := C - a$ for some fixed $a \in C$. Clearly $\tilde{C}$ is still convex, open, and nonempty. If we find x^* for $\tilde{C}$ and $x_0 - a$ such that

$$x^*(x) < x^*[x_0 - a] \quad \forall x \in \tilde{C}$$

then

$$x^*(x + a) < x^*[x_0] \quad \forall x \in \tilde{C} = C - a$$

or equivalently

$$x^*(z) < \tilde{x}^*[x_0] \quad \forall z \in C.$$

So, from now on assume that $0 \in C$.

We introduce the *Minkowski functional*, $m_C : X \to \mathbb{R}$

$$m_C(x) := \inf \left\{ t > 0 : \quad \frac{1}{t} x \in C \right\}.$$

Since C is open and $0 \in C$ for each $x \in X$ there exists $t > 0$ such that $\frac{1}{t} x \in C$. Thus $m_C(x) < \infty$ for all $x \in X$.

Also we observe the following

$$\frac{1}{t} x \in C \quad \forall t > m_C(x). \tag{10.4}$$

Indeed let $t > m_C(x)$. By the definition of the infimum there exists $t_0 \in [m_C(x), t)$ such that $\frac{1}{t_0} x \in C$. Since C is convex and $0 \in C$ we find that then also

$$\frac{1}{t} x = \frac{t_0}{t} \frac{1}{t_0} x + \left(1 - \frac{t_0}{t}\right) 0 \in C.$$

This establishes (10.4)

Also, there exist $\Lambda > 0$ such that

$$m_C(x) \leq \Lambda \|x\|_X. \tag{10.5}$$

Indeed, since $0 \in C$ and C is open, there exists $\delta > 0$ such that $B(0, \delta) \in C$ and thus for each $x \in X$ we have $\frac{\delta x}{2\|x\|} \in B(0, \delta) \in C$. Thus, $m_C(x) \leq \frac{2}{\delta} \|x\|$.

Now we claim that the Minkowski functional is sublinear. It is easy to see that for $\lambda > 0$

$$\begin{aligned} m_C(\lambda x) =& \inf\left\{t > 0 : \quad \frac{1}{t}\lambda x \in C\right\} \\ =& \lambda \inf\left\{\frac{t}{\lambda} > 0 : \quad \frac{\lambda}{t}x \in C\right\} \\ =& \lambda \inf\left\{\tilde{t} > 0 : \quad \tilde{t}x \in C\right\} \\ =& \lambda m_C(x). \end{aligned}$$

Now assume that $x, y \in C$.

Fix any $t > m_C(x)$ and $s > m_C(y)$. Then we have $\frac{1}{s}x \in C$ and $\frac{1}{t}y \in C$, by (10.4).

We then have by convexity of C,

$$\frac{1}{t+s}(x+y) = \underbrace{\frac{t}{t+s}}_{\in[0,1]}(\frac{1}{t}x) + \underbrace{\frac{s}{t+s}}_{=1-\frac{t}{t+s}}(\frac{1}{s}y) \in C$$

Thus,

$$m_C(x+y) \leq t+s.$$

This holds for any $t > m_C(x)$ and any $s > m_C(y)$. Letting $t \to m_C(x)^+$ and $s \to m_C(y)^+$ we conclude that m_C is indeed sublinear.

Next we observe that m_C seperates C and x_0, in the sense that

$$m_C(x_0) \geq 1, \quad \text{and} \quad m_C(x) < 1 \quad \forall x \in C.$$

Indeed, $m_C(x) < 1$ for all $x \in C$ follows from the fact that $\frac{1}{1}x \in C$ and C is open, so there must be some $t < 1$ such that $\frac{1}{t}x \in C$ as well – and thus $m_C(x) \leq t < 1$.

To see $m_C(x_0) \geq 1$ observe that if $m_C(x_0) < 1$ then from (10.4) we know $x_0 = 1x_0 \in C$ which is a contradiction. Thus $m_C(x_0) \geq 1$ (observe it could indeed be $= 1$ if x_0 lies on the boundary of C!).

Now set $Y := \operatorname{span}\{x_0\}$ and define $y^* \in Y^*$ as

$$y^*[\lambda x_0] := \lambda m_C(x_0), \quad \lambda \in \mathbb{R}.$$

Since $m_C(x) \geq 0$ we find that

$$y^*[\lambda x_0] \begin{cases} = m_C(\lambda x_0) & \text{if } \lambda > 0 \\ \leq 0 \leq m_C(\lambda x_0) & \text{if } \lambda \leq 0 \end{cases}$$

Thus $y^*[y] \leq m_C(y)$ for all $y \in Y$. By the Hahn-Banach theorem we find $x^* : X \to \mathbb{R}$ with

$$\begin{gathered} x^* : X \to \mathbb{R} \quad \text{linear, and} \\ x^*[x] \leq m_C(x) \quad \forall x \in X. \end{gathered}$$

In particular, in view of (10.5) we have that $x^* \in X^*$.

Thus we have

$$x^*[x] \leq m_C(x) < 1 \quad \forall x \in C$$

and

$$x^*[x_0] = y^*[x_0] = m_C(x_0) \geq 1.$$

□

One can also separate two disjoint convex sets

Theorem 10.22. *Let X be a normed space, $U, V \subset X$ convex and nonempty. Assume that U is open, then there exists $\lambda \in \mathbb{R}$ and a functional $x^* \in X^*$ such that*

$$x^*[u] < \lambda \leq x^*[v] \quad \forall u \in U\,, v \in V.$$

Proof. We set

$$C := U - V := \{u - v : \quad u \in U, v \in V\}.$$

Since U is open, so is C. Since U and V are convex, so is C. Since $U \cap V = \emptyset$ we have $0 \notin C$. By Theorem 10.21 there exist $x^* \in X^*$ such that

$$x^*[x] < 0 \quad \forall x \in C.$$

That is,

$$x^*[u - v] < 0 \quad \forall u \in U, v \in V.$$

That is

$$x^*[u] < x^*[v] \quad \forall u \in U, v \in V.$$

Fixing $u \in U$ we see that $\lambda := \inf_{v \in V} x^*[v] \in \mathbb{R}$, and we have

$$x^*[u] \leq \lambda \leq x^*[v] \quad \forall u \in U, v \in V.$$

We need to make the first $\leq$ into a $<$.

Assume to the contrary, that there exists some $u \in U$ such that $x^*[u] = \lambda$. We know that $x^* \neq 0$ (because we have the strict inequality above), so there exists some vector $p \in X$ such that $x^*[p] > 0$. Since U is open there exists some $\delta > 0$ such that $u + \delta p \in U$, and thus

$$x^*[\underbrace{u + \delta p}_{\in U}] > x^*[u] = \lambda,$$

which is a contradiction. So indeed we have

$$x^*[u] < \lambda \leq x^*[v] \quad \forall u \in U, v \in V.$$

□

Theorem 10.23 (Strict separation theorem)**.** *Let X be a normed space, $C \subset X$ be a nonempty, closed, convex set, and let $x_0 \in X \setminus C$.*

Then there exists $x^* \in X^*$ and $\lambda \in \mathbb{R}$ such that

$$x^*[c] \leq \lambda < x^*[x_0] \quad \forall c \in C.$$

Proof. C is closed so $X \setminus C$ is open, and since $x_0 \in X \setminus C$ there exists a small ball $B(x_0, \delta) \subset X \setminus C$. Apply Theorem 10.22 to $B(x_0, \delta)$ and C, we find $x^* \in X^*$ and $\lambda \in \mathbb{R}$ such that

$$x^*[x] < \lambda \leq x^*[c] \quad \forall c \in C, x \in B(x_0, \delta).$$

Multiplying this with (-1) we find for $y^* := -x^*$

$$y^*[c] \leq -\lambda < y^*[x_0],$$

which is what we wanted. $\square$

11. The bidual and reflexivity

We can define the dual X^*, and Hahn-Banach tells us that X^* tells us a lot about X. So why not discuss X^{**}, the bidual. We first observe the following

Every element of $x \in X$ can be identified as an element in X^{**} by the following procedure

For $x \in X$ we define $x^{**} \in X^{**}$ via

$$x^{**}[y^*] := y^*[x] \quad \text{for } y^* \in X^*.$$

Clearly

$$|x^{**}[y^*]| \leq \|x\|_X \, \|y^*\|_{X^*},$$

so $\|x^{**}\|_{X^{**}} \leq \|x\|_X$. On the other hand, from Hahn-Banach, Corollary 10.14,

$$\|x\|_X = \sup_{y^* \in X^*, \|y^*\|=1} |y^*[x]| = \sup_{y^* \in X^*, \|y^*\|=1} |x^{**}[y^*]| \leq \|x^{**}\|_{X^{**}}.$$

That is, we have $\|x^{**}\|_{X^{**}} = \|x\|_X$.

We denote the map $x \mapsto x^{**}$ by $J_X : X \to X^{**}$ and call it the *canonical embedding of $X^{**} \hookrightarrow X$*. Clearly $J_X : X \to X^{**}$ is linear and continuous, and it is an isometry $\|J_X x\|_{X^{**}} = \|x\|_X$ (in particular it is injective).

Theorem 11.1. *$J_X : X \to X^{**}$ is linear, injective, and an isometry.*

Example 11.2. Let $p \in (1, \infty)$ and take any $f^{**} \in (L^p(\mathbb{R}^n))^{**}$. This is a functional acting on $g^* \in (L^p(\mathbb{R}^n))^*$. By the Riesz representation theorem, Theorem 5.39, for any g^* there exists $g \in L^{p'}(\mathbb{R}^n)$ such that $\|g\|_{L^{p'}} = \|g^*\|_{(L^p)^*}$ and

$$g^*(h) = \int_{\mathbb{R}^n} hg \quad \forall h \in L^p(\mathbb{R}^n).$$

There is a one-to-one relationship between g^* and g. So f^{**} induces a functional f^* on $L^{p'}$ in the following way

$$f^*[g] := f^{**}[g^*] \quad g \in L^{p'}(\mathbb{R}^n).$$

But then f^* is a linear functional of $L^{p'}$ so there exist $f \in L^p(\mathbb{R}^n)$ such that for all g,

$$\int fg = f^*[g] = f^{**}[g^*].$$

Now let us consider the canonical embedding $J_X f$. We have by definition,

$$J_X f[g^*] = g^*[f] = \int_{\mathbb{R}^n} gf = f^*[g] = f^{**}[g^*].$$

That is $J_X f = f^{**}$. That is $J_X : L^p(\mathbb{R}^n) \to (L^p(\mathbb{R}^n))^{**}$ is surjective!

This is a very nice property (as we shall see), so we give it a name: *reflexivity*.

Observe our argument above fails for $p = \infty$ and $p = 1$, since we cannot apply Riesz representation theorem for L^∞! The reason is below, Lemma 11.12

Definition 11.3. Let X be a normed vector space. If $J_X : X \to X^{**}$ is *surjective*, then we say that X is *reflexive*.

Remark 11.4.
- We often say that reflexivity means $X^{**} = X$. This is dangerous (still we'll do it), because equality is not really defined. So it is important to remind ourselves now and then that the equality must be under the canonical mapping J_X.

Exercise 11.5. *Let X be reflexive. Show that X is necessarily complete.*

Hint: X^{**} is always complete (it is a dual space!). What happens to $J_X(x_n)$ if $(x_n)_n$ is a Cauchy sequence?

Exercise 11.6. *Let X be reflexive, then X^* is reflexive.*

We will sharpen Exercise 11.6 for Banach spaces, see Theorem 11.11

Exercise 11.7. *Let X and Y be isomorphic. I.e. assume that there exists $T : X \to Y$ linear and bounded and bijective, and $T^{-1} : Y \to X$ is linear bounded. Then X is reflexive if and only if Y is reflexive.*

In particular conclude that if X is equipped with two equivalent norms $\|\cdot\|_1$ and $\|\cdot\|_2$, then $(X, \|\cdot\|_1)$ is reflexive if and only if $(X, \|\cdot\|_2)$ is reflexive.

Exercise 11.8. *Assume that a normed space $(X, \|\cdot\|_X)$ can be written as $X = X_1 \times X_2$ where $(X_i, \|\cdot\|_{X_i})$ are two normed spaces. Assume furthermore that the norm $\|\cdot\|_X$ is equivalent to*

$$X = X_1 \times X_2 \ni (x_1, x_2) \mapsto \|x_1\|_{X_1} + \|x_2\|_{X_2}.$$

Show that if X_1 and X_2 are reflexive, so is X.

Theorem 11.9. *Let X be a reflexive Banach space and let $U \subset X$ be a closed subspace. Then U is reflexive as well.*

Proof. Let $u^{**} \in U^{**}$. We argue similar to Example 10.1: any functional $x^* \in X^*$ can be considered as an element of $x^*\Big|_U \in U^*$. So set

$$x^{**}(x^*) := u^{**}[x^*\Big|_U],$$

then $x^{**} \in X^{**}$.

Since X is reflexive, there exists $x \in X$ such that $x^{**} = J_X x$.

First, we show that $\underline{x \in U}$. Assume this is not the case. Since U is closed by assumption, we can apply Corollary 10.18 and find $x^* \in X^*$ such that $x^*(u) = 0$ for all $u \in U$ but $x^*(x) = 1$.

But then

$$x^*[x] = J_X x[x^*] = x^{**}[x^*] = u^{**}[\underbrace{x^*\Big|_U}_{\equiv 0}] = 0.$$

This is a contradiction, so $x \in U$.

It remains to $\underline{\text{show that } J_U x = u^{**}}$. Fix $u^* \in U^*$, then by Hahn-Banach Corollary 10.9 there exists an extension $x^* \in X^*$ with $x^*\Big|_U = u^*$.

Then

$$J_U x[u^*] = u^*[x] \overset{x \in U}{=} x^*[x] = J_X x[x^*] = x^{**}[x^*] = u^{**}[x^*\Big|_U] = u^{**}[u^*].$$

We can conclude. □

Corollary 11.10. *$W^{1,p}$ is reflexive.*

Proof. It is clear that $W^{1,p} \subset L^p$ w.r.t L^p-norm, but $W^{1,p}$ is not closed under the L^p-norm!

So we rather use the identifaction used for the dual, see the proof of Corollary 10.12. In that sense $W^{1,p}(\mathbb{R}^n)$ is a closed subspace of $L^p(\mathbb{R}^n) \times \ldots L^p(\mathbb{R}^n)$ which is reflexive (cf. Exercise 11.8, Footnote 28). □

The following sharpens Exercise 11.6 for Banach spaces

Theorem 11.11. *A Banach space X is reflexive if and only if X^* is reflexive.*

Proof. We already have shown that if X is reflexive then so is X^*, Exercise 11.6.

So assume X^* is reflexive. Then we know that X^{**} is reflexive. Since X is complete and J_X is an isometry, $J_X(X) \subset X^{**}$ is a closed subspace. Thus, by Theorem 11.9, $J_X(X)$ is also reflexive. But $J_X(X)$ and X are (by definition) isometric isomorphic, so by Exercise 11.7 we conclude that X must be be reflexive. □

Lemma 11.12. *$L^1(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n)$ are not reflexive.*

Proof. By Exercise 11.6 and the Riesz Representation theorem (Which identifies $(L^1)^*$ with L^∞) it suffices to that the dual L^∞ is not L^1, namely there exists a functional $T \in (L^\infty(\mathbb{R}^n))^*$ which cannot be represented as a integration agains an L^1-function.

Set

$$Tf := f(0).$$

This is a linear functional on $C^0 \cap L^\infty(\mathbb{R}^n)$. It is also bounded. By Hahn-Banach, Corollary 10.9, there exists an extension map $T^e : L^\infty(\mathbb{R}^n) \to L^\infty(\mathbb{R}^n)$.

Now assume that $T = J_X f$ for some $f \in L^1(\mathbb{R}^n)$. Then for any $g \in (L^1(\mathbb{R}^n))^* = L^\infty(\mathbb{R}^n)$ we'd have

$$J_X f[g] = \int_{\mathbb{R}^n} fg$$

In particular for continuous and bounded functions g we'd have

$$\int_{\mathbb{R}^n} fg = g(0).$$

But this means that $f = \delta_0$ in the sense of distributions, which can't be because $fd\mathcal{L}^n$ is absolutely continuous with respect to the Lebesgue measure whereas δ_0 is not.

Thus $(L^1)^{**} = (L^\infty)^* \supsetneq L^1$, which means that L^1 is not reflexive.[29]. By Theorem 11.11 this implies that L^1 can also not be reflexive. $\square$

Exercise 11.13. *Let X be a finite dimensional normed vector space, then X is reflexive.*

12. Weak Convergence & Reflexivity

The main goal of this section is to reap the fruits of reflexivity: a weak version of Bolzano-Weierstrass theorem.

Theorem 12.1 (Bounded sets are weakly precompact (in reflexive spaces))**.** *Assume X be a reflexive space. Then every bounded sequence $(x_n)_{n\in\mathbb{N}}$ has a weakly convergent subsequence.*

Clearly we will need to define what weakly convergent subsequence means. Theorem 12.1 is incredibly important, it is often references as "*by reflexivity*". More precisely (albeit still incorrect) it is the *Eberlein–Smulian Theorem* or (worse, because thats about weak*-convergence, which implies this theorem: *Banach-Alaoglu Theorem*). A slightly better version of refering to Theorem 12.1 is weak compactness (in reflexive spaces).

We will prove this theorem later, the proof is a bit lengthy, at the end of the section. More important than the proof are the applications (for once)

Let us define weak and weak*-convergence.

Definition 12.2 (Weak convergence)**.** Let X be a normed vector space and X^* its dual.

[29] Observe that we needed Hahn-Banach, i.e. the axiom of choice. This is necessary, there is no explicit functional on L^∞ that is not in L^1

(1) Let $\{x_n\}_{n\in\mathbb{N}} \subset X$ a sequence. We say that $(x_n)_{n\in\mathbb{N}}$ *weakly converges* to x in X,
$$x_n \rightharpoonup x$$
if
$$x^*[x_n] \xrightarrow{n\to\infty} x^*[x].$$
(2) Let $\{x_n^*\}_{n\in\mathbb{N}} \subset X^*$ a sequence. We say that $(x_n^*)_{n\in\mathbb{N}}$ *weakly$*$ converges* to x^* in X^*, if
$$x_n^*[x] \xrightarrow{n\to\infty} x^*[x] \quad \forall x \in X.$$
In particular, if X is reflexive, weak*-convergence is the same as the weak convergence in X^*.
(3) when we want to emphasize the contrast, we refer to the usual X-convergence as *convergence in norm* or *strong convergence*. I.e. Let $\{x_n\}_{n\in\mathbb{N}} \subset X$ a sequence. We say that $(x_n)_{n\in\mathbb{N}}$ *strongly converges* to x in X if $\|x_n - x\|_X \xrightarrow{n\to\infty} 0$.

Remark 12.3. Having defined (sequentially) weak convergence, we naturally obtain a notion of weakly closed sets $A \subset X$ is *weakly closed* if and only if any weakly converging sequence a_n has its weak limit in A if all $a_n \in A$. Thus we can define open sets by the complent of closed sets. Thus weak convergence introduces a topology.

However, unless X is finite dimensional, there is no metric inducing this topology.

12.1. **Basic Properties of weak convergence.** Alright, so now need to cover the basics for weak convergence in general

Lemma 12.4. *Weak limits are unique.*

Proof. Assume x_k converges weakly to x, and at the same time x_k weakly converges to y. Then for any $x^* \in X^*$
$$x^*[x] \xleftarrow{k\to\infty} x^*[x_k] \xrightarrow{k\to\infty} x^*[y]$$
That is
$$x^*[x] = x^*[y],$$
so
$$x^*[x-y] = 0.$$
This holds for any x^*, so by Exercise 10.15, $x - y = 0$, that is $x = y$. □

Exercise 12.5. *Show that weak*-limits are unique*

Exercise 12.6.
- *Strong convergence implies weak convergence*
- *If X is reflexive, weak convergence in X^* and weak*-convergence in X^* are the same.*

Lemma 12.7. *(1) The norm is* lower semicontinuous *under weak convergence. That is, assume x_k weakly converge to x. Then*
$$\|x\|_X \le \liminf_{k\to\infty} \|x_k\|_X$$

(2) *the dual norm is lower semicontinuous under weak* convergence. That is, assume x_k^* weak*-converges to x^*. Then*
$$\|x^*\|_{X^*} \leq \liminf_{k\to\infty} \|x_k^*\|_{X^*}$$

Proof. (1) Assume x_k weakly converge to x. By Hahn-Banach, Corollary 10.14, there exists $x^* \in X^*$, $\|x^*\|_{X^*} = 1$, such that
$$\|x\|_X = x^*[x] = \lim_{k\to\infty} x^*[x_k] \leq \underbrace{\|x^*\|_{X^*}}_{\leq 1} \liminf_{k\to\infty} \|x_k\|_X.$$

(2) Assume x_k^* weak*-converges to x^*. Let $\varepsilon > 0$. There exist $x \in X$, $\|x\|_X \leq 1$ such that
$$\|x^*\|_X \leq x^*[x] + \varepsilon = \lim_{k\to\infty} x_k^*[x] + \varepsilon \leq \liminf_{k\to\infty} \|x_k^*\|_{X^*} \underbrace{\|x\|_X}_{=1} + \varepsilon.$$
Letting $\varepsilon \to 0$ we conclude.

□

The typical example of weak convergence is usually given in ℓ^2-spaces, see Example 12.24; Let for $i \in \mathbb{N}$
$$e_i = (0,\ldots,0, \underbrace{1}_{i\text{-th position}}, 0, \ldots)$$
We will argue that e_i weakly converges to 0 in ℓ^2, showing that the inequality in Lemma 12.7 can indeed be strict. For L^p-spaces see Example 12.12.

12.2. **Weak convergence in L^p-spaces.** it is important to *observe that "weak L^p"-convergence has* **nothing** *to to with "weak L^p"-space from Definition 3.49.*

As usual the fundamental examples are L^p and $W^{1,p}$.

Example 12.8. • For $p \in [1,\infty)$ let $f_k \in L^p(\mathbb{R}^n)$. By Riesz representation, any linear functional $g^* \in (L^p(\mathbb{R}^n))^*$ can be identified with
$$g^*[f_k] = \int_{\mathbb{R}^n} f_k g,$$
where $g \in L^{p'}(\mathbb{R}^n)$. Thus f_k weakly converges to f in $L^p(\mathbb{R}^n)$ if and only if
$$\int_{\mathbb{R}^n} (f_k - f)g = 0 \quad \forall g \in L^{p'}(\mathbb{R}^n)$$

Usually we use test-functions (i.e. $C_c^\infty(\mathbb{R}^n)$).

Exercise 12.9. *Let $p \in (1,\infty)$.*

(1) *Show that f_k weakly converges to f in $L^p(\mathbb{R}^n)$ if and only if*
- $\sup_k \|f_k\|_{L^p(\mathbb{R}^n)} + \|f\|_{L^p(\mathbb{R}^n)} < \infty$, *and*

- $\lim_{k\to\infty}\int_{\mathbb{R}^n}(f_k - f)\varphi = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$.

(2) The boundedness assumption above is necessary. Namely, give an example of $(f_k)_{k\in\mathbb{N}} \subset L^p(\mathbb{R}^n)$ and $f \in L^p(\mathbb{R}^n)$ with

$$\lim_{k\to\infty}\int_{\mathbb{R}^n}(f_k - f)\varphi = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

but where f_k does not weakly converge to f in $L^p(\mathbb{R}^n)$.

Remark 12.10. So in some sense, weak convergence is very similar to "pointwise convergence" for distributions. Namely if f_k converges (say) L^p-weakly to f, then if we think of f_k as a distribution

$$f_k[\varphi] \xrightarrow{k\to\infty} f[\varphi] \quad \forall \varphi \in \mathscr{S}(\mathbb{R}^n).$$

Lemma 12.11. *Let $p \in (1,\infty)$. Let $f_k \in L^p(\mathbb{R}^n)$ and assume that $\sup_{k\in\mathbb{N}} \|f_k\|_{L^p(\mathbb{R}^n)} < \infty$. Moreover assume that $f_k(x) \to 0$ for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$. Then f_k weakly converges to zero in $L^p(\mathbb{R}^n)$.*

Proof of Lemma 12.11. Set

$$\Lambda := \sup_k \|f_k\|_{L^p} < \infty.$$

Let $\varphi \in C_c^\infty(\mathbb{R}^n)$. Since φ has compact support we can apply Egorov in $\operatorname{supp}\varphi$, Theorem 3.38, and find for any ε a compact set $K \subset \mathbb{R}^n$ such that $\|f_k - 0\|_{L^\infty} \xrightarrow{k\to\infty} 0$ and $\mathcal{L}^n(\operatorname{supp}\varphi \setminus K) < \varepsilon$. Then by Hölder's inequality

$$\left|\int_{\mathbb{R}^n} f_k\varphi\right| \le \|f_k\|_{L^\infty(K)}\|\varphi\|_{L^1(\mathbb{R}^n)} + \|f_k\|_{L^1(\operatorname{supp}\varphi\setminus K)}\,\|\varphi\|_{L^\infty(\mathbb{R}^n)}.$$

Since $p > 1$,

$$\|f_k\|_{L^1(\operatorname{supp}\varphi\setminus K)} \le \mathcal{L}^n(\operatorname{supp}\varphi \setminus K)^{1-\frac{1}{p}}\,\|f_k\|_{L^p} \le \mathcal{L}^n(\operatorname{supp}\varphi \setminus K)^{1-\frac{1}{p}}\,\Lambda \le \varepsilon\Lambda.$$

So we have shown

$$\left|\int_{\mathbb{R}^n} f_k\varphi\right| \le \|f_k\|_{L^\infty(K)}\|\varphi\|_{L^1(\mathbb{R}^n)} + \Lambda\varepsilon \xrightarrow{k\to\infty} \Lambda\varepsilon.$$

This holds for any $\varepsilon > 0$, so

$$\lim_{k\to\infty}\left|\int_{\mathbb{R}^n} f_k\varphi\right| = 0 \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n). \tag{12.1}$$

□

Example 12.12. Let $p > 1$. Pick any $f \in C_c^\infty(\mathbb{R}^n)$ with $\|f\|_{L^p(\mathbb{R}^n)} > 0$. Set

$$f_k(x) := k^{\frac{n}{p}} f(kx).$$

Then $\|f_k\|_{L^p(\mathbb{R}^n)} = \|f\|_{L^p(\mathbb{R}^n)}$, that is $\sup_{k\in\mathbb{N}} \|f_k\|_{L^p(\mathbb{R}^n)} < \infty$.

Observe that since $f(kx) = 0$ whenever $x \neq 0$ and $k \gg 1$ (depending on x) so we have

$$f_k(x) \xrightarrow{k\to\infty} 0 \quad \forall x \neq 0.$$

That is, $f_k(x) \to 0$ a.e. in $\mathbb{R}^n$. Since $p > 1$ we can use Lemma 12.11 to conclude that f_k weakly converges to zero in $L^p(\mathbb{R}^n)$.

Now let $q < p$, then we have

$$\|f_k\|_{L^q(\mathbb{R}^n)} = k^{\frac{n}{p}-\frac{n}{q}}\|f\|_{L^q(\mathbb{R}^n)} \leq k^{\frac{n}{p}-\frac{n}{q}}C(\operatorname{supp} f)\|f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

That is, f_k strongly converges to zero in $L^q(\mathbb{R}^n)$ for any $q \in [1,p)$. But f_k converges only weakly to zero in $L^p(\mathbb{R}^n)$.

We record a reformulation of Theorem 12.1 for $L^p(\Omega)$-spaces

Theorem 12.13. *Let $\Omega \subset \mathbb{R}^n$ open, $1 \leq p < \infty$. Assume $f_k \in L^p(\Omega)$ with $\sup_{k\in\mathbb{N}} \|f_k\|_{L^p(\Omega)} < \infty$. Then there exists a subsequence f_{k_i} and some $f \in L^p(\Omega)$ such that f_{k_i} weakly converges to f in $L^p(\Omega)$, i.e.*

$$\int_\Omega f_{k_i}\varphi \xrightarrow{k\to\infty} \int_\Omega f\varphi \quad \forall \varphi \in L^{p'}(\Omega).$$

Moreover we have

$$\|f\|_{L^p(\Omega)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(\Omega)}.$$

Proof. Since we can extend $f \in L^p(\Omega)$ to $\mathbb{R}^n$ by setting $\tilde{f} := \chi_\Omega f$ this follows easily from the $\mathbb{R}^n$-theorem. The estimate follows from Lemma 12.7. □

Exercise 12.14. *Let $\Omega \subset \mathbb{R}^n$ open. Let $(f_n)_{n\in\mathbb{N}} \subset L^1(\Omega)$ and $f \in L^1(\Omega)$. Show that if f_n weakly converges to f in $L^1(\Omega)$ then*

$$\int_\Omega (f_n - f)\varphi \xrightarrow{n\to\infty} 0$$

for each $\varphi \in C_c^\infty(\Omega)$.

Exercise 12.15 (Weak-Strong Products)**.** *Assume $p, q, r \in (1,\infty)$ such that $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$.*

Let $f_n, f \in L^p(\mathbb{R}^n)$, $g_n, g \in L^q(\mathbb{R}^n)$ and assume that f_n converges weakly to f in $L^p(\mathbb{R}^n)$ and g_n converges strongly to g in $L^q(\mathbb{R}^n)$.

Show that $f_n g_n$ converges weakly to fg in $L^r(\mathbb{R}^n)$.

Products of weakly convergent sequences may not converge weakly without additional assumptions – which lead e.g. to the *div-curl*-lemma which details such assumptions.

12.3. Weak convergence in Sobolev space.

Proposition 12.16. *Let $1 \leq p < \infty$. Assume $(f_k)_{k\in\mathbb{N}} \subset W^{1,p}(\mathbb{R}^n)$ and $f \in L^p(\mathbb{R}^n)$.*

Then f_k converges weakly to f in $W^{1,p}(\mathbb{R}^n)$ if and only if

1. *f_k converges weakly to f in $L^p(\mathbb{R}^n)$*
2. *$\partial_\alpha f_k$ converges weakly in L^p to some $F_\alpha \in L^p(\mathbb{R}^n)$, $\alpha = 1, \ldots, n$.*

In both cases $f \in W^{1,p}(\mathbb{R}^n)$ *and* $F_\alpha = \partial_\alpha f$.

Proof. By Corollary 10.12 each $T \in (W^{1,p}(\mathbb{R}^n))^*$ can be represented as

$$T[f] = \int_{\mathbb{R}^n} gf + \int_{\mathbb{R}^n} G \cdot Df,$$

where $g, G \in L^{p'}(\mathbb{R}^n)$.

Thus weak convergence in $W^{1,p}(\mathbb{R}^n)$ is equivalent to

$$\int_{\mathbb{R}^n} g(f_k - f) + \int_{\mathbb{R}^n} G \cdot D(f_k - f) \xrightarrow{k\to\infty} 0 \quad \forall g, G \in L^{p'}. \tag{12.2}$$

$\underline{\Rightarrow}$ Assume f_k converges f weakly in $W^{1,p}$. Choose $G = 0$ in (12.2) to get (1). Choose $g = 0$ and $G = (0, \ldots, 0, \tilde{g}, 0, \ldots, 0)$ to get (2).

$\underline{\Leftarrow}$ So let us assume (1) and (2). We clearly get

$$T[f_k - f] = \int_{\mathbb{R}^n} g(f_k - f) + \int_{\mathbb{R}^n} G \cdot (Df_k - F) = \int_{\mathbb{R}^n} g(f_k - f) + \sum_{\alpha=1}^n \int_{\mathbb{R}^n} G_\alpha(\partial_\alpha f_k - F_\alpha) \xrightarrow{k\to\infty} 0.$$

However, who is to tell us that $f \in W^{1,p}(\mathbb{R}^n)$?

Well let $g = 0$ and choose $G := (0, 0, \ldots, \varphi, 0 \ldots)$ (where $\varphi \in C_c^\infty(\mathbb{R}^n)$ is in the α-position). We then have

$$\int_{\mathbb{R}^n} \varphi F_\alpha = \lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi \partial_\alpha f_k.$$

On the other hand, by the definition of weak derivative,

$$\int_{\mathbb{R}^n} \varphi \partial_\alpha f_k = -\int_{\mathbb{R}^n} \partial_\alpha \varphi f_k$$

So we have

$$\int_{\mathbb{R}^n} \varphi \partial_\alpha f = -\lim_{k\to\infty} \int_{\mathbb{R}^n} \partial_\alpha \varphi f_k.$$

But now, by (1) we have

$$\lim_{k\to\infty} \int_{\mathbb{R}^n} \underbrace{\partial_\alpha \varphi}_{\in L^{p'}} f_k = \int_{\mathbb{R}^n} \partial_\alpha \varphi f.$$

So we have shown that

$$\int_{\mathbb{R}^n} \varphi F_\alpha = \lim_{k\to\infty} \int_{\mathbb{R}^n} \varphi \partial_\alpha f_k = \int_{\mathbb{R}^n} \partial_\alpha \varphi f.$$

That is $\partial_\alpha f = F_\alpha \in L^p(\mathbb{R}^n)$ (in distributional sense), and in view of Theorem 4.46 we conclude that $f \in W^{1,p}(\mathbb{R}^n)$. $\square$

12.4. More involved basic properties of weak convergence.

Theorem 12.17. *Let X be a Banach space. Then weak and weak* convergent sequences are necessarily bounded.*

So Theorem 12.1 is kind of the converse of Theorem 12.17.

For the proof of Theorem 12.17 we need the

Theorem 12.18 (Banach-Steinhaus or Uniform Boundedness Principle)**.** *Let X be a Banach space and Y a normed vector space. Suppose that $\mathcal{F}$ is a (possibly uncountable) family of continuous, linear operators $T \in L(X,Y)$. If $\mathcal{F}$ is pointwise bounded, that is if for all $x \in X$ we have*

$$\sup_{T\in\mathcal{F}} \|Tx\| < \infty$$

then $\mathcal{F}$ is uniformly bounded in norm, i.e.

$$\sup_{T\in\mathcal{F}} \|T\|$$

Usually one uses the *Baire category theorem* to prove this statement, but one can avoid this and give an elementary proof. We follow [Sokal, 2011].

Lemma 12.19. *Let T be a bounded linear operator from the normed spaces X to Y. Then for any $x \in X$ and any $r > 0$ we have*

$$\sup_{y\in B(x,r)} \|Ty\| \geq \|T\|\, r.$$

here $B(x,r) = \{y \in X : \ \|y-x\| < r\}$ is the open r-ball.

Proof. Fix $x \in X$ and $r > 0$. Let $z \in X$ then $z = \frac{1}{2}(z+x) + \frac{1}{2}(z-x)$ so

$$\|Tz\| \leq \frac{1}{2}\left(\|T(x+z)\| + \|T(x-z)\|\right) \leq \max\{\|T(x+z)\|, \|T(x-z)\|\}.$$

Consequently,

$$\|T\| = \sup_{\|\tilde{z}\|<1} \|T\tilde{z}\| = \frac{1}{r}\sup_{\|z\|<r} \|Tz\| \leq \frac{1}{r}\sup_{y\in B(x,r)} \|T(y)\|.$$

□

Proof of Theorem 12.18. Suppose to the contrary that

$$\sup_{T\in\mathcal{F}} \|T\| = \infty.$$

Then there must be $(T_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ with $\|T_n\| \geq 4^n$.

Set $x_0 = 0$. Apply Lemma 12.19 to x_0 and $r = 3^{-1}$. Then there must be $x_1 \in X$, with

$$\|x_1 - x_0\|_X < 3^{-1}$$

but
$$\|T_1x_1\| \geq \frac{2}{3}3^{-1}\,\|T_1\|.$$
Repeating this inductively we find $x_n \in X$ such that $\|x_n - x_{n-1}\|_X \leq 3^{-n}$ and
$$\|T_nx_n\| \geq \frac{2}{3}3^{-n}\|T_n\|.$$
In particular x_n is a Cauchy sequence in X, since
$$\|x_n - x_m\| \leq \sum_{k=\min\{n,m\}+1}^{\max\{n,m\}} \|x_k - x_{k-1}\| \leq \sum_{k=\min\{n,m\}+1}^{\max\{n,m\}-1} 3^{-k} \xrightarrow{\min\{n,m\}\to\infty} 0.$$
Since X is complete there exists x such that $\lim_{n\to\infty} x_n = x$, and we have
$$\|x_n - x\| \leq \sum_{k=n+1}^{\infty} 3^{-k} = \frac{3}{2}3^{-n-1} = \frac{1}{2}3^{-n}.$$
Then we have
$$\|T_nx\| \geq \|T_nx_n\| - \|T_n(x-x_n)\| \geq \frac{2}{3}3^{-n}\|T_n\| - \frac{1}{2}3^{-n}\|T_n\| = \frac{1}{6}3^{-n}\|T_n\| \geq \frac{1}{6}\left(\frac{4}{3}\right)^n.$$
That is,
$$\sup_{n\in\mathbb{N}} \|T_nx\|_X = +\infty,$$
a contradiction to the assumption. □

Proof of Theorem 12.17 – weak convergence. Let $(x_n)_{n\in\mathbb{N}}$ be a weakly convergent sequence to $x \in X$.

Let $J_X : X \to X^{**}$ be the canonical embedding, then weak convergence implies
$$(J_Xx_n)[x^*] = x^*[x_n] \xrightarrow{n\to\infty} x^*[x] \quad \text{in } \mathbb{R}$$
Since convergent sequences in $\mathbb{R}$ are bounded, we find that for any $x^* \in X^*$
$$\sup_{n\in\mathbb{N}} |(J_Xx_n)[x^*]| < \infty.$$
Since X^* is a Banach space we can apply Banach-Steinhaus Theorem 12.18 and find that actually
$$\sup_{n\in\mathbb{N}} \|J_Xx_n\|_{X^{**}} < \infty,$$
which by Theorem 11.1 implies that
$$\sup_{n\in\mathbb{N}} \|x_n\|_X < \infty.$$
□

Exercise 12.20. *Prove Theorem 12.17 for weak*-convergence.*

A corollary of Theorem 12.17 is the following, which also implies that Theorem 12.1 is an honest extension of Bolzano-Weierstrass theorem for finite dimensional sets ("bounded sets are pre-compact") to infinite dimensional sets ("bounded sets are *weakly* precompact").

Exercise 12.21. *If X is finite dimensional, then weak convergence coincides with strong convergence.*

We also obtain

Lemma 12.22 (Pointwise a.e. and weak L^p-limit coincide)**.** *Let $p \in (1,\infty)$. Assume $f_k \in L^p(\mathbb{R}^n)$ weakly converges to some $f \in L^p(\mathbb{R}^n)$. Assume that moreover $f_k(x) \to g(x)$ for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$. Then $f = g \in L^p(\mathbb{R}^n)$.*

Proof. g as a pointwise limit of measurable functions is measurable and from Fatou's lemma Corollary 3.9 applied to $|f_k|^p$ we have

$$\|g\|_{L^p(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)}.$$

The right-hand side is finite by Theorem 12.17, so $g \in L^p(\mathbb{R}^n)$. Now we can argue similar to the proof of Lemma 12.11 to conclude. □

12.5. **Applications of weak compactness theorem - Theorem 12.1.** Let us discuss some applications of Theorem 12.1.

Theorem 12.23. *Let X be reflexive, and $x^* \in X^*$. Then there exists $x \in X$, $\|x\| = 1$ such that*

$$x^*[x] = \|x^*\|_{X^*}$$

Proof. If $x^* = 0$ then there is nothing to show. So assume $\|x^*\|_{X^*} > 0$. We have

$$\|x^*\|_{X^*} = \sup_{\|x\|\leq 1} x^*[x].$$

So let $x_k \in X$ such that

$$x^*[x_k] \xrightarrow{k\to\infty} \|x^*\|_{X^*}.$$

Since $\|x_k\|_X \leq 1$, by Theorem 12.1 we can pass to a subsequence (relabel if necessary) and have that x_k weakly converges to $x \in X$. By weak lower semicontinuity of the norm, Lemma 12.7, we have $\|x\| \leq 1$. On the other hand we have

$$x^*[x] = \lim_{k\to\infty} x^*[x_k] = \|x^*\|_{X^*}.$$

It remains to show that $\|x\|_X = 1$. Set $\lambda := \|x\|_X$. Then

$$\frac{1}{\lambda}\|x^*\|_{X^*} = \frac{1}{\lambda}x^*[x] = x^*[x/\lambda] \leq \|x^*\|_{X^*} \underbrace{\|x/\lambda\|_X}_{=1} = \|x^*\|_{X^*}$$

Dividing both sides by $\|x^*\|_{X^*} > 0$ we find that $\frac{1}{\lambda} \leq 1$, i.e. $\lambda = 1$. □

Example 12.24. The unit sphere in ℓ^2 is a typical example against Bolzano-Weierstrass. Take

$$e_i = (0, \ldots, 0, 1, 0, \ldots) \in \ell^2,$$

then $\|e_i\|_{\ell^2} = 1$, so $(e_i)_{i\in\mathbb{N}}$ is uniformly bounded. However $\|e_i - e_j\|_{\ell_2} = \sqrt{2}$ so there is no subsequence of $(e_i)_i$ that is strongly convergent.

However a subsequence e_i weakly converges by Theorem 12.1. What is the limit? It is zero. Indeed the dual space for $\ell^2(\mathbb{N})$ is $\ell^2(\mathbb{N})$ in the sense that any element $T \in (\ell^2(\mathbb{N}))^*$ corresponds to some $(c_k)_{k\in\mathbb{N}} \in \ell^2$ such that

$$T[f] = \sum_k c_k f_k \quad \forall f \in \ell^2.$$

Since $c_k \in \ell^2(\mathbb{N})$ we have $\lim_{k\to\infty} c_k = 0$, so

$$T[e_i] = c_i \xrightarrow{i\to\infty} 0.$$

This holds for any $T \in (\ell^2)^*$, so e_i weakly converges to zero!

In particular the usual example, Example 12.24, show that the unit sphere $\{x \in \|x\| : \|x\|_X = 1\}$ albeit closed and bounded, might not be *weakly closed*

Definition 12.25. A set $A \subset X$ is called *weakly closed* if any weakly convergent sequence $(a_n) \subset A$, $a_n \xrightarrow{*} x \in X$ has its limit in A, i.e. $x \in A$.

Since strong convergence implies weak convergence, Exercise 12.6, any weakly closed set is also closed. The reverse does not hold (see above) but we have

Theorem 12.26. *Let X be a normed space and $Y \subset X$ be convex. Then Y is weakly closed if and only if Y is closed.*

Proof. Since strong convergence implies weak convergence, Exercise 12.6, any weakly closed set is closed.

So now assume that Y is closed and convex and consider a weakly convergent sequence $y_k \in Y$, $y_k \rightharpoonup x \in X$. Assume $x \notin Y$. By the strict separation theorem, Theorem 10.23, there exists $x^* \in X^*$ and $\lambda \in \mathbb{R}$ such that

$$x^*[x_n] \leq \lambda < x^*[x] \quad \forall n \in \mathbb{N}$$

But then $\lim_{n\to\infty} x^*[x_n] \neq x^*[x]$, contradiction. □

Exercise 12.27. *Let X be a reflexive Banach space and $(x_n)_{n\in\mathbb{N}} \subset X$ a weak Cauchy sequence, i.e. for all $x^* \in X^*$ we have $(x^*[x_n])_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Show that $(x_n)_{n\in\mathbb{N}}$ converges weakly.*

Exercise 12.28 (Mazur's theorem)**.** *Let X be a normed vector space and $(x_n)_{n\in\mathbb{N}}$ a sequence with $x_n \overset{*}{\to} x$. Show that there exists a sequence $(y_n)_{n\in\mathbb{N}}$ of convex combinations*

$$y_n = \sum_{k=1}^{N_n} \lambda_{n,k} x_k, \quad \textit{with } \sum_{k=1}^{N_n} \lambda_{n,k} = 1,\ \lambda_{n,k} \in [0,1],\ N_n \in \mathbb{N}$$

such that y_n converges (strongly!) to x.

Hint: Consider the convex hull

$$C := \left\{ \sum_{k=1}^{N} \lambda_k x_k : \quad \sum_{k=1}^{N} \lambda_k = 1, \lambda_k \in [0,1], N \in \mathbb{N} \right\}.$$

Exercise 12.29. *Let X and Y be Banach spaces, and $(x_n)_{n\in\mathbb{N}} \subset X$ with $x_n \rightharpoonup x \in X$. Let $T \in L(X,Y)$. Show that $Tx_n \rightharpoonup Tx$ in Y.*

12.6. **Application: Direct Method of Calculus of Variations & Tonelli's theorem.** On particularly important example is the following *energy method* or *direct method of the Calculus of Variations* for Partial Differential Equations.

Theorem 12.30. *Let $f \in L^2(\mathbb{R}^n)$, $\lambda > 0$, then there exists a unique $u \in W^{1,2}(\mathbb{R}^n)$ such that*

$$\Delta u - \lambda u = f \quad \textit{in } \mathbb{R}^n$$

holds in distributional sense, where $\Delta = \sum_{i=1}^n \partial_{x_i}\partial_{x_i}$, i.e.

$$-\int_{\mathbb{R}^n} \nabla u \cdot \nabla\varphi - \lambda \int u\varphi = \int_{\mathbb{R}^n} f\varphi$$

Proof. Define the *energy*

$$E(u) := \frac{1}{2}\int_{\mathbb{R}^n} |Du|^2 + \frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu.$$

For any $u \in W^{1,2}(\mathbb{R}^n)$ the energy is finite $E(u) < \infty$.

The energy is also *coercive* in $W^{1,2}(\mathbb{R}^n)$. This means any energy bounded sequence with $(u_k)_k \in W^{1,2}(\mathbb{R}^n)$, with $\sup_k E(u_k) < \infty$ also satisfies $\sup_k \|u_k\|_{W^{1,2}(\mathbb{R}^n)} < \infty$. Indeed, from Hölder's inequality and Young's inequality, for any $\varepsilon > 0$,

$$\begin{aligned}
\frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu \geq& \frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 - \|f\|_{L^2(\mathbb{R}^n)} \|u\|_{L^2(\mathbb{R}^n)} \\
=& \frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 - \frac{1}{\varepsilon}\|f\|_{L^2(\mathbb{R}^n)}\, \varepsilon\|u\|_{L^2(\mathbb{R}^n)} \\
\geq& \frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 - \varepsilon^2\|u\|^2_{L^2(\mathbb{R}^n)} - \frac{1}{\varepsilon^2}\|f\|^2_{L^2(\mathbb{R}^n)} \\
=& (\frac{\lambda}{2} - \varepsilon^2)\|u\|^2_{L^2(\mathbb{R}^n)} - \frac{1}{\varepsilon^2}\|f\|^2_{L^2(\mathbb{R}^n)}.
\end{aligned}$$

Taking $\varepsilon^2 < \frac{\lambda}{4}$ we find

$$\frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu \geq c_\lambda \|u\|^2_{L^2(\mathbb{R}^n)} - C_\lambda \|f\|^2_{L^2(\mathbb{R}^n)}.$$

Thus

$$E(u) \geq c_\lambda \|u\|^2_{W^{1,2}(\mathbb{R}^n)} - C_\lambda \|f\|^2_{L^2(\mathbb{R}^n)},$$

i.e.

$$\|u\|^2_{W^{1,2}(\mathbb{R}^n)} \leq \tilde{C}\left(E(u) + \|f\|^2_{L^2(\mathbb{R}^n)}\right).$$

In particular whenever $\sup_k E(u_k) < \infty$ also $\sup_k \|u_k\|_{W^{1,2}(\mathbb{R}^n)} < \infty$. That is, E is coercive in $W^{1,2}(\mathbb{R}^n)$.

Now we can apply the the *direct method* of Calculus of Variation to minimize E in $W^{1,2}(\mathbb{R}^n)$.

Set

$$I := \inf_{u \in W^{1,2}(\mathbb{R}^n)} E(u).$$

We see that $I \leq 0$ since $I \leq E(0) = 0$. We also have $I > -\infty$, since as before with Young and Hoelder inequality

$$\frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu \geq (\frac{\lambda}{2} - \varepsilon^2)\|u\|^2_{L^2} - \frac{1}{\varepsilon^2}\|f\|^2_{L^2(\mathbb{R}^n)},$$

so if $\varepsilon^2 < \frac{\lambda}{2}$ we have

$$E(u) \geq \frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu \geq -\frac{1}{\varepsilon^2}\|f\|^2_{L^2(\mathbb{R}^n)} > -\infty.$$

That is I is a finite number.

By the definition of the infimum there must be a sequence $(u_k)_{k\in\mathbb{N}} \subset W^{1,2}(\mathbb{R}^n)$ such that

$$\lim_{k\to\infty} E(u_k) = I.$$

In particular we then have $\sup_k |E(u_k)| < \infty$, and thus by coercivity $\|u_k\|_{W^{1,2}(\mathbb{R}^n)}$. By Corollary 11.10 $W^{1,2}$ is reflexive, so by Theorem 12.1 (up to passing to a subsequence) we can assume that u_k converges to some $u \in W^{1,2}(\mathbb{R}^n)$ weakly in $W^{1,2}(\mathbb{R}^n)$. By Proposition 12.16 we have that ∇u_k weakly converges to ∇u in L^2, and u_k converges weakly to u in L^2. By *lower semicontinuity* of the L^2-norm, Lemma 12.7, we have

$$\int_{\mathbb{R}^n} |\nabla u|^2 \leq \liminf_{k\to\infty} \int_{\mathbb{R}^n} |\nabla u_k|^2$$

and

$$\int_{\mathbb{R}^n} |u|^2 \leq \liminf_{k\to\infty} \int_{\mathbb{R}^n} |u_k|^2.$$

By weak L^2-convergence we also have in view of Example 12.8,

$$\int_{\mathbb{R}^n} uf = \lim_{k\to\infty} \int_{\mathbb{R}^n} u_k f.$$

So we find

$$E(u) \leq \liminf_{k\to\infty} E(u_k) = I.$$

On the other hand $u \in W^{1,2}$ so
$$E(u) \geq I.$$
This means $E(u) = I$, that is u is a minimizer of E in $W^{1,2}(\mathbb{R}^n)$.

The next step is to show that u satisfies an equation, called the *Euler-Lagrange equation*. This is Fermat's theorem: if $x \in \mathbb{R}^n$ minimizes a smooth F then $F'(x) = 0$. Here F' becomes the *first variation* of E – the proof is the same.

More precisely, since u is minimizer we have for any $\varphi \in C_c^\infty(\mathbb{R}^n)$
$$E(u) \leq E(u + t\varphi) \quad \forall t \in \mathbb{R}.$$
So we have
$$0 \leq \liminf_{t \to 0} \frac{E(u+t\varphi) - E(u)}{t}.$$
Let us look at the right-hand site.
$$\int_{\mathbb{R}^n} |\nabla(u+t\varphi)|^2 - \int_{\mathbb{R}^n} |\nabla u|^2 = 2t \int_{\mathbb{R}^n} \nabla u \cdot \nabla \varphi.$$
$$\int_{\mathbb{R}^n} |(u+t\varphi)|^2 - \int_{\mathbb{R}^n} |u|^2 = 2t \int_{\mathbb{R}^n} u\varphi.$$
and
$$\int_{\mathbb{R}^n} (u+t\varphi) f - \int_{\mathbb{R}^n} uf = t \int_{\mathbb{R}^n} \varphi f$$
That is, for each $t \neq 0$,
$$\frac{E(u+t\varphi) - E(u)}{t} = \int_{\mathbb{R}^n} \nabla u \cdot \nabla \varphi + \int_{\mathbb{R}^n} u\varphi + \int_{\mathbb{R}^n} \varphi f.$$
So we have found that
$$0 \leq \int_{\mathbb{R}^n} \nabla u \cdot \nabla \varphi + \int_{\mathbb{R}^n} u\varphi + \int_{\mathbb{R}^n} \varphi f \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$
Switching φ by $-\varphi$ we obtain
$$0 = \int_{\mathbb{R}^n} \nabla u \cdot \nabla \varphi + \lambda \int_{\mathbb{R}^n} u\varphi + \int_{\mathbb{R}^n} \varphi f \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$
That is u solves the equation that we wanted it to solve.

Lastly we need to show uniqueness. Assume that $u, v \in W^{1,2}$ both solve
$$0 = \int_{\mathbb{R}^n} \nabla u \cdot \nabla \varphi + \lambda \int_{\mathbb{R}^n} u\varphi + \int_{\mathbb{R}^n} \varphi f \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$
and
$$0 = \int_{\mathbb{R}^n} \nabla v \cdot \nabla \varphi + \lambda \int_{\mathbb{R}^n} v\varphi + \int_{\mathbb{R}^n} \varphi f \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$
This equation is linear in u: we can subtract the second equation from the first and obtain for $w := u - v$
$$0 = \int_{\mathbb{R}^n} \nabla w \cdot \nabla \varphi + \lambda \int_{\mathbb{R}^n} w\varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

Now plug in $\varphi = w$ (this is ok, since we can approximate w by $w_k \in C_c^\infty(\mathbb{R}^n)$, i.e. *by density*).

$$\|\nabla w\|_{L^2}^2 + \lambda \|w\|_{L^2}^2 = 0.$$

Since $\lambda > 0$ this implies $\|w\|_{L^2} = 0$ i.e. $w = 0$ a.e., i.e. $u = v$ a.e. This proves uniqueness. □

Observe $\lambda > 0$ was important here, the same statement may not be true for $\lambda < 0$.

The direct method needs coercivity and lower semicontinuity, so it is not too difficult to copy the above proof to obtain

Theorem 12.31 (Tonelli)**.** *Let X be a reflexive normed vector space and $f : X \to \mathbb{R}$ is weakly lower semicontinuous, i.e.*

$$f(x) \leq \liminf_{k\to\infty} f(x_k) \quad \text{whenever } x_k \text{ weakly converges to } x.$$

Assume $U \subset X$ is nonempty and one of the following holds

- *U is weakly closed and $f : U \to \mathbb{R}$ is coercive, that is whenever $(u_k) \subset U$ is a sequence such that $\|u_k\|_X \to \infty$ then $f(u_k) \to \infty$.*
- *U is bounded, convex and closed*

Then there exist $\bar{u} \in U$ such that

$$f(\bar{u}) = \min_{u\in U} f(u).$$

Exercise 12.32. *Prove Theorem 12.31.*

12.7. **Proof of Theorem 12.1.** The first step is in the proof of Theorem 12.1 is to work with the X^* and weak*-convergence.

Theorem 12.33 (Banach-Alaoglu)**.** *Let X be a normed space and separable. Then any sequence $(x_k^*)_k \subset X^*$ with $\sup_k \|x_k^*\|_k < \infty$ has a weak*-convergent subsequence.*

Observe that for linear functionals $\sup_k \|x_k^*\|_k < \infty$ implies equicontinuity. If X was a *compact* metric space we could try to argue by Arzela-Ascoli. Indeed, in the proof of Arzela-Ascoli, compactness is used for some sort of separability – so since X is separable, we will use the ideas of the proof of Arzela-Ascoli.

Proof of Theorem 12.33. By renormalizing (i.e. othweise considering x_k^*/K for $K := \sup \|x_k^*\|$) we can assume that

$$\sup_k \|x_k^*\|_k \leq 1.$$

Since X is separable, we can find a countable dense subset of X, let us denote it by $\{x_n : n \in \mathbb{N}\}$.

Then for each fixed $n \in \mathbb{N}$

$$\sup_k |x_k^*[x_n]| \leq \|x_n\|,$$

that is $(x_k^*[x_n])_{k\in\mathbb{N}} \subset \mathbb{R}$ is bounded, so there exists a converging subsequence $x_{k_{i;n}}^*[x_n])_{i\in\mathbb{N}} \subset \mathbb{R}$. Taking a diagonal sequence we find a subsequence $(x_{k_i}^*)_{i\in\mathbb{N}}$ sich that for each $n \in \mathbb{N}$ the sequence $(x_{k_i}^*[x_n])_{i\in\mathbb{N}} \subset \mathbb{R}$ converges. We denote its limit

$$x^*[x_n] := \lim_{i\to\infty} x_{k_i}^*[x_n].$$

Let now

$$Z := \operatorname{span}(\{x_n\}) \equiv \{z = \sum_{j=1}^N \lambda_j x_j \quad N \in \mathbb{N}, \lambda_j \in \mathbb{R}\}.$$

This is a linear space, and by the linearity of $x_{k_i}^*$ we can extend x^* to a linear functional on Z,

$$x^*[\sum_{j=1}^N \lambda_j x_j] := \sum_{j=1}^N \lambda_j \lim_{i\to\infty} x_{k_i}^*[x_j] \equiv \lim_{i\to\infty} x_{k_i}^*[\sum_{j=1}^N \lambda_j x_j].$$

We then have

$$x^*[\sum_{j=1}^N \lambda_j x_j] \leq \underbrace{\limsup_{i\to\infty} \|x_{k_i}^*\|_{X^*}}_{\leq 1} \|\sum_{j=1}^N \lambda_j x_j\|_X,$$

that is $x^* \in Z^*$ with $\|x^*\|_{Z^*} \leq 1$. Since x^* is uniformly continuous on Z and $Z \subset X$ is dense, we can extend x^* uniquely to all of X and find a linear functional $x^* \in X^*$.

Now let $x \in X$ and $\varepsilon > 0$. There exists $n \in \mathbb{N}$ such that $\|x - x_n\|_X < \varepsilon$. Then

$$|x_{k_i}^*(x) - x^*(x)| \leq 2\varepsilon + |x_{k_i}^*(x_n) - x^*(x_n)|$$

So,

$$\limsup_{i\to\infty} |x_{k_i}^*(x) - x^*(x)| \leq 2\varepsilon.$$

This holds for any $\varepsilon > 0$, so we have

$$\limsup_{i\to\infty} |x_{k_i}^*(x) - x^*(x)| = 0.$$

That is $x_{k_i}^*(x) \xrightarrow{i\to\infty} x^*(x)$ for all $x \in X$. That is $x_{k_i}^*$ weak*-converges to x^*. □

Exercise 12.34. *Show that without separability of X the statement of Theorem 12.33 may fail. Consider for example $e_n^* \in (\ell^\infty(\mathbb{N}))^*$ given by*

$$e_n^*[x] := x_n \quad x \in \ell^\infty.$$

Show that there is no weak-convergent subsequence.*

The statement of Theorem 12.1, called the *Theorem of Eberlein-Smulian* is then a consequence of Theorem 12.33, using that weak* convergence for $(X^*)^*$ is the same as weak convergence in X. We can get rid of the separability, because we only need to work in the closure of the space spanned by the sequence – by definition a separable space.

Proof of Theorem 12.1. Let $(x_n)_{n\in\mathbb{N}}$ be such that $\sup_n \|x_n\|_X < \infty$.

Set

$$Y := \overline{\operatorname{span}\{x_n\}} \equiv \overline{\{z = \sum_{j=1}^N \lambda_j x_j \quad N \in \mathbb{N}, \lambda_j \in \mathbb{R}\}}.$$

Observe that Y is a closed subspace of X, and thus by Theorem 11.9, Y is also reflexive. Moreover Y is separable, thus by reflexivity $Y^{**} \cong Y$ is separable, and thus by Corollary 10.20, Y^* is separable.

Let $J_Y : Y \to Y^{**}$ be the canonical embedding, then $x_n^{**} := J_Y x \in Y^{**}$ is a bounded sequence in $(Y^*)^*$, so by Theorem 12.33 there exists a weak*-convergent subsequence $(x_{n_i}^{**})_{i\in\mathbb{N}}$ to some $x^{**} \in Y^{**}$. Since Y is reflexive, there exists exactly one $x \in X$ such that $J_Y x = x^{**}$.

Let $x^* \in X^*$ then we have

$$x^*[x_{n_i}] = x_{n_i}^{**}[x^*] \xrightarrow{i\to\infty} x^{**}[x^*] = x^*[x].$$

That is x_{n_i} weakly converges to x. □

12.8. **Weak convergence and compactness for L^1 and Radon measures.** Theorem 12.1 can not be applied to $L^1(\mathbb{R}^n)$ because $L^1(\mathbb{R}^n)$ is not reflexive.

Example 12.35. Let $\eta \in C_c^\infty(\mathbb{R}^n)$, $\int_{\mathbb{R}^n} \eta = 1$, $\eta(0) = 1$. Also assume for simplicity that $\eta(-x) = \eta(x)$.

Set

$$\eta_k(x) := k^{-n}\eta(kx).$$

Then

$$\|\eta_k\|_{L^1(\mathbb{R}^n)} = \|\eta\|_{L^1(\mathbb{R}^n)} < \infty,$$

that is $(\eta_k)_{k\in\mathbb{N}}$ is uniformly bounded in $L^1(\mathbb{R}^n)$. Let now $\varphi \in C_c^\infty(\mathbb{R}^n)$, then we have (here we use the symmetry of η)

$$\int_{\mathbb{R}^n} \eta_k(y)\, \varphi(y) dy = \eta_k * \varphi(0) \xrightarrow{k\to\infty} \varphi(0).$$

That is η_k "weakly converges" to the measure δ_0, in the sense that

$$\int_{\mathbb{R}^n} \varphi(y)\, \eta_k(y) dy \xrightarrow{k\to\infty} \int_{\mathbb{R}^n} \varphi(y) d\delta_0(y) \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

So we do have a weak convergence, just the space $L^1(\mathbb{R}^n)$ is not "really weakly closed".

Any $f \in L^1(\mathbb{R}^n)$ can be considered as a Radon measure $f\llcorner\mathcal{L}^n$. If $\sup_k \|f_k\|_{L^1(\mathbb{R}^n)} < \infty$ we have

$$\sup_k f_n\llcorner\mathcal{L}^n(K) < \infty \quad \forall \text{compact } K.$$

It turns out that this is the right notion in which Theorem 12.1 indeed works.

Definition 12.36. Let $\mu, (\mu_k)_{k=1}^{\infty}$ be Radon measures on $\mathbb{R}^n$. We say that μ converges weakly to the measure μ (in the sense of Radon measures), $\mu_k \rightharpoonup \mu$ if one of the following statements is satisfied

(1) For all $f \in C_c^0(\mathbb{R}^n)$ we have
$$\int_{\mathbb{R}^n} f d\mu_k \xrightarrow{k\to\infty} \int_{\mathbb{R}^n} f d\mu$$
(2) $\limsup_{k\to\infty} \mu_k(K) \leq \mu(K)$ for all compact sets $K \subset \mathbb{R}^n$ and $\mu(U) \leq \liminf_{k\to\infty} \mu_k(U)$ for each open set $U \subset \mathbb{R}^n$

(3) $\lim_{k\to\infty} \mu_k(B) = \mu(B)$ for each bounded Borel set $B \subset \mathbb{R}^n$ with $\mu(\partial B) = 0$.

Lemma 12.37. *The three conditions in Definition 12.36 are equivalent.*

Proof. $\underline{(1) \Rightarrow (2)}$ Assume (1) holds. Fix $K \subset \mathbb{R}^n$ be compact and $U \supset K$ open. Then we can find $f \in C^0(\mathbb{R}^n)$, $f \equiv 0$ in $\mathbb{R}^n \backslash U$ and $f \equiv 1$ in K. Indeed, we have $\varepsilon := \operatorname{dist}(K, \mathbb{R}^n \backslash U) > 0$ (exercise), so if we set
$$f(x) := \max\{1 - \frac{1}{\varepsilon}\operatorname{dist}(x, K), 0\}$$
we see that f is as required. Then we have by (1)
$$\mu(K) \leq \int_{\mathbb{R}^n} f d\mu = \lim_{k\to\infty} \int_{\mathbb{R}^n} f d\mu_k \leq \liminf_{k\to\infty} \mu_k(U).$$
and
$$\limsup_{k\to\infty} \mu_k(K) \leq \limsup_{k\to\infty} \int_{\mathbb{R}^n} f d\mu_k = \int_{\mathbb{R}^n} f d\mu \leq \mu(U).$$
The claim now follows from Theorem 1.68, because
$$\mu(U) = \sup_{K\subset U,\quad K \text{ compact}} \mu(K) \leq \liminf_{k\to\infty} \mu_k(U),$$
and
$$\mu(K) = \inf_{K\subset U,\ U \text{ open}} \mu(U) \geq \limsup_{k\to\infty} \mu_k(K).$$

$\underline{(2) \Rightarrow (3)}$ Assume (2) holds. Let $B \subset \mathbb{R}^n$ be a bounded Borel set with $\mu(\partial B) = 0$. Then
$$\begin{aligned}\mu(B) =& \mu(\underbrace{B \setminus \partial B}_{\text{open}}) \overset{(2)}{\leq} \liminf_{k\to\infty} \mu_k(B \setminus \partial B)\\ &\leq \limsup_{k\to\infty} \mu_k(\underbrace{\overline{B}}_{\text{compact}}) \overset{(2)}{\leq} \mu(\overline{B}) = \mu(\overline{B} \setminus \partial B) \leq \mu(B).\end{aligned}$$
Thus,
$$\mu(B) = \liminf_{k\to\infty} \mu_k(B \setminus \partial B) \leq \liminf_{k\to\infty} \mu_k(B)$$
and
$$\mu(B) = \limsup_{k\to\infty} \mu_k(\overline{B}) \geq \limsup_{k\to\infty} \mu_k(B)$$

which readily gives

$$\mu(B) = \lim_{k\to\infty} \mu_k(B)$$

(3) $\Rightarrow$ (1) Assume (3) holds and let $f \in C_c^0(\mathbb{R}^n)$. W.l.o.g. we can assume that $f \geq 0$ everywhere (split $f = f_+ - f_-$ otherwise). We can also assume that $\|f\|_{L^\infty} \leq 1$ (otherwise divide by $\|f\|_{L^\infty}$).

Take a large open ball B such that $\operatorname{supp} f \subset B$ and moreover $\mu(\partial B) = 0$. Observe that while not every ball B must satisfy $\mu(\partial B) = 0$, by Proposition 1.72 we can find such a ball.

Fix $\varepsilon > 0$, and pick $N \approx \frac{1}{\varepsilon}$ many t_i,

$$0 = t_0 < t_1 < \ldots < t_N \leq 2$$

such that $|t_i - t_{i-1}| < \varepsilon$ and $t_N \geq \|f\|_{L^\infty}$ – and we may also assume $\mu(f^{-1}(t_i)) = 0$, whenever $i \geq 1$. This is possible by Exercise 1.74. Since $\|f\|_{L^\infty} \leq 1$ we then have

$$\sum_{i=1}^N t_{i-1}\mu(f^{-1}[t_{i-1}, t_i] \cap \overline{B}) \leq \int_{\mathbb{R}^n} f d\mu \leq \sum_{i=1}^N t_i\mu(f^{-1}[t_{i-1}, t_i] \cap \overline{B})$$

and

$$\sum_{i=1}^N t_{i-1}\mu_k(f^{-1}[t_{i-1}, t_i] \cap \overline{B}) \leq \int_{\mathbb{R}^n} f d\mu_k \leq \sum_{i=1}^N t_i\mu_k(f^{-1}[t_{i-1}, t_i] \cap \overline{B})$$

Since f is continuous, $f^{-1}[t_{i-1}, t_i]$ is closed and thus $B_i := f^{-1}[t_{i-1}, t_i] \cap \overline{B}$ is compact and in particular a bounded Borel set. Also

$$\mu(\partial B_i) \leq \mu(\partial\overline{B}) + \mu(\{t_{i-1}\}) + \mu(\{t_i\}).$$

Subtracting both inequalities we find

$$\int_{\mathbb{R}^n} f d\mu - \int_{\mathbb{R}^n} f d\mu_k \leq \sum_{i=1}^N t_i\mu(f^{-1}[t_{i-1}, t_i] \cap \overline{B}) - \sum_{i=1}^N t_{i-1}\mu_k(f^{-1}[t_{i-1}, t_i] \cap \overline{B})$$

and

$$\int_{\mathbb{R}^n} f d\mu_k - \int_{\mathbb{R}^n} f d\mu \leq \sum_{i=1}^N t_i\mu_k(f^{-1}[t_{i-1}, t_i] \cap \overline{B}) - \sum_{i=1}^N t_{i-1}\mu(f^{-1}[t_{i-1}, t_i] \cap \overline{B})$$

Taking the limit, using condition (2), we find that

$$\begin{aligned}
\limsup_{k\to\infty} \left|\int_{\mathbb{R}^n} f d\mu - \int_{\mathbb{R}^n} f d\mu_k\right| \leq& \sum_{i=1}^N (t_i - t_{i-1})\, \mu(f^{-1}[t_{i-1}, t_i] \cap \overline{B})\\
\leq& \varepsilon \sum_{i=1}^N \mu(f^{-1}[t_{i-1}, t_i] \cap \overline{B})\\
\leq& 2\varepsilon\mu(\overline{B})
\end{aligned}$$

This holds for all $\varepsilon > 0$ so we have shown that

$$\limsup_{k\to\infty} \left|\int_{\mathbb{R}^n} f d\mu - \int_{\mathbb{R}^n} f d\mu_k\right| = 0.$$

That is, (1), is established. $\square$

Exercise 12.38. *Show that we can equivalently change (1) in Definition 12.36 into*

For all $f \in C_c^\infty(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} f d\mu_k \xrightarrow{k\to\infty} \int_{\mathbb{R}^n} f d\mu$$

Theorem 12.39 (Weak compactness for measures)**.** *Let $(\mu_k)_{k=1}^\infty$ be a sequence of Radon measures on $\mathbb{R}^n$ satisfying for any compact $K \subset \mathbb{R}^n$,*

$$\sup_k \mu_k(K) < \infty$$

Then there exists a subsequence $(\mu_{k_i})_{i\in\mathbb{N}}$ and a Radon measure μ such that

$$\mu_{k_i} \rightharpoonup \mu$$

in the sense of Radon measures, Definition 12.36.

We will skip the proof, but record a consequence.

Exercise 12.40. *If $f_k \in L^1(\mathbb{R}^n)$ with*

$$\sup_k \|f_k\|_{L^1(\mathbb{R}^n)} < \infty$$

then there exists a subsequence $(f_{k_i})_{i\in\mathbb{N}}$ and two Radon measures μ_+ and μ_- such that

$$\int_{\mathbb{R}^n} f_{k_i}\varphi \xrightarrow{i\to\infty} \int_{\mathbb{R}^n} \varphi d\mu_+ - \int_{\mathbb{R}^n} \varphi d\mu_-.$$

12.9. Yet another definition of L^p and $W^{1,p}$.

In Proposition 4.40 and Definition 4.42 we reinterpreted (and defined) the L^p-space and $W^{1,p}$-space as metric completion, i.e. we said for $1\leq p < \infty$

$$f \in L^p(\mathbb{R}^n) \quad \text{iff there exists } f_k \in C^\infty(\mathbb{R}^n): \quad \|f_k\|_{L^p(\mathbb{R}^n)} < \infty \text{ and } \|f_k - f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

and

$$f \in W^{1,p}(\mathbb{R}^n) \quad \text{iff there exists } f_k \in C^\infty(\mathbb{R}^n)(\mathbb{R}^n): \quad \|f_k\|_{W^{1,p}(\mathbb{R}^n)} < \infty \text{ and } \|f_k - f\|_{W^{1,p}(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

(Ok, we actually did not have C^∞ there, but C_c^∞ – but these two notions are equivalent by looking at $f_k\eta_{B(0,k)}$ instead of f_k).

Now for $1<p < \infty$ we can obtain the following characterizations (observe this way one really eliminates the need for Lebesgue integrals, in comparison to Riemann integrals)

Proposition 12.41. *Let $1<p < \infty$ and $f : \mathbb{R}^n \to \mathbb{R}$. Then $f \in L^p(\mathbb{R}^n)$ if and only if there exists $f_k \in C^\infty(\mathbb{R}^n)$ with $\sup_k \|f_k\|_{L^p(\mathbb{R}^n)} < \infty$ and $f_k(x) \xrightarrow{k\to\infty} f(x)$ for almost every $x \in \mathbb{R}^n$.*

Proof. If $f \in L^p(\mathbb{R}^n)$ there exist $f_k \in C_c^\infty(\mathbb{R}^n)$ converging strongly to f, in particular $\sup_k \|f_k\|_{L^p} < \infty$ – and up to a subsequence f_k converges a.e. to f.

For the other direction observe that if $f_k \in C^\infty(\mathbb{R}^n)$ with $\sup_k \|f_k\|_{L^p(\mathbb{R}^n)} < \infty$ by reflexivity Theorem 12.1 there exists $g \in L^p(\mathbb{R}^n)$ and a subsequence $f_{k_i} \rightharpoonup g$ in $L^p(\mathbb{R}^n)$, that is

$$\int_{\mathbb{R}^n} f_{k_i}\varphi \xrightarrow{i\to\infty} \int_{\mathbb{R}^n} g\varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

In view of Lemma 12.22 $g = f$ since f_k converges a.e. to f and $p > 1$. □

Similarly one can define

Exercise 12.42. *Let* $1<p<\infty$ *and* $f : \mathbb{R}^n \to \mathbb{R}$. *Then* $f \in W^{1,p}(\mathbb{R}^n)$ *if and only if there exists* $f_k \in C^\infty(\mathbb{R}^n)$ *with* $\sup_k \|f_k\|_{W^{1,p}(\mathbb{R}^n)} < \infty$ *and* $f_k(x) \xrightarrow{k\to\infty} f(x)$ *for almost every* $x \in \mathbb{R}^n$.

The above does not work for $p = 1$, the space that comes out is not $W^{1,1}$, but BV, cf. Section 15

12.10. **Compact operators.** Recall the definition of compact operator and compact embedding from Definition 8.11.

Later we will see, that if Ω is a smoothly bounded set (e.g. a ball) then $W^{1,p}(\Omega)$ embeds compactly in $L^p(\Omega)$, $p \in [1,\infty]$ – this is called the *Rellich-Kondrachov Theorem* Theorem 13.35. This can be used for example when we use the direct method to solve PDEs as in Theorem 12.30 – but with *lower order nonlinerarity*, e.g. $\Delta u - \lambda u + |u|^2 u = f$.

Compact operators T take weakly convergent sequences $(x_k)_k$ into strongly convergent sequences $(Tx_k)_{k\in\mathbb{N}}$.

Exercise 12.43. *Let* $(X, \|\cdot\|)$ *and* $(Y, \|\cdot\|)$ *and* $T \in \mathcal{L}(X,Y)$ *a linear bounded operator which moreover is compact.*

Assume that $(x_k)_{k\in\mathbb{N}}$ *are weakly convergent in* X. *Show that* $(Tx_k)_{k\in\mathbb{N}}$ *is strongly convergent.*

Hint: Show that $(Tx_k)_{k\in\mathbb{N}}$ *is weakly convergent by establishing that* $y^*[T\cdot] \in X^*$ *for* $y^* \in Y^*$. *Then use that weak and strong limit coincide.*

13. Sobolev spaces

A remark on literature: A standard reference for Sobolev spaces is [Adams and Fournier, 2003]. Very readable is also [Evans and Gariepy, 2015]. The introduction here takes a lot from the introduction to Sobolev spaces in [Evans, 2010]. A classical reference Sobolev spaces in PDEs is [Gilbarg and Trudinger, 2001]. Also [Ziemer, 1989]. For very delicate problems one might also consult [Maz'ya, 2011].

We now define the notion of Sobolev space on $\Omega \subset \mathbb{R}^n$ (where we always will think of Ω as an open set). For $\Omega = \mathbb{R}^n$ and $p \in [1,\infty)$ we know this is the same as our old definition by approximation, cf. Theorem 4.46. Observe in the next definition $p = \infty$ is included.

Definition 13.1. (1) Let $1 \leq p \leq \infty$, $k \in \mathbb{N}$ and $\Omega \subset \mathbb{R}^n$ open, nonempty. The *Sobolev space* $W^{k,p}(\Omega)$ is the set of functions

$$u \in L^p(\Omega)$$

such that for any multiinidex γ, $|\gamma| \leq k$ we find a function (the *distributional γ-derivative* or *weak γ-derivative*) *"$\partial^\gamma u$"*$\in L^p(\Omega)$ such that

$$\int_\Omega u\,\partial^\gamma \varphi = (-1)^{|\gamma|} \int_\Omega \text{"}\partial^\gamma u\text{"}\,\varphi \quad \forall \varphi \in C_c^\infty(\Omega).$$

Such u are also sometimes called Sobolev-functions.

(2) For simplicity we write $W^{0,p} = L^p$.

(3) The norm of the Sobolev space $W^{k,p}(\Omega)$ is given as

$$\|u\|_{W^{k,p}(\Omega)} = \sum_{|\gamma| \leq k} \|\partial^\gamma u\|_{L^p(\mathbb{R}^n)}$$

or equivalently (exercise!)

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\gamma| \leq k} \|\partial^\gamma u\|^p_{L^p(\mathbb{R}^n)} \right)^{\frac{1}{p}}.$$

(4) We define another Sobolev space $H^{k,p}(\Omega)$ as follows

$$H^{k,p}(\Omega) = \overline{C^\infty(\overline{\Omega})}^{\|\cdot\|_{W^{k,p}(\Omega)}}.$$

that is the (metric) closure or completion of the space $(C^\infty(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$. In yet other words, $H^{k,p}(\Omega)$ consists of such functions $u \in L^p(\Omega)$ such that there exist approximations $u_k \in C^\infty(\overline{\Omega})$ with

$$\|u_k - u\|_{W^{k,p}(\Omega)} \xrightarrow{k\to\infty} 0.$$

We will later see that $H^{k,p}$ is the same as $W^{k,p}$ locally, or for nice enough domains; and use the notation H or W interchangeably. For $k = 0$ this fact follows from Lemma 4.38 for any open set Ω.

(5) Now we introduce the Sobolev space $H_0^{k,p}(\Omega)$

$$H_0^{k,p}(\Omega) = \overline{C_c^\infty(\overline{\Omega})}^{\|\cdot\|_{W^{k,p}(\Omega)}}.$$

We will later see that this space consists of all maps $u \in H^{k,p}(\Omega)$ that satisfy $u, \nabla u, \ldots \nabla^{k-1} u \equiv 0$ on $\partial\Omega$ in a suitable sense (the *trace sense*, for a precise formulation see Theorem 13.31). – Again, later we see that $H = W$ and thus, $W_0^{k,p}(\Omega) = H_0^{k,p}(\Omega)$ for nice sets Ω.

Observe that in view of Lemma 4.38, $L^p(\Omega) = W^{0,p}(\Omega) = W_0^{0,p}(\Omega)$.

(6) The local space $W^{k,p}_{loc}(\Omega)$ is similarly defined as $L^p_{loc}(\Omega)$. A map belongs to $u \in W^{k,p}_{loc}(\Omega)$ if for any $\Omega' \subset\subset \Omega$ we have $u \in W^{k,p}(\Omega')$.

Remark 13.2. Some people write $H^{k,p}(\Omega)$ instead of $W^{k,p}(\Omega)$. Other people use $H^k(\Omega)$ for $H^{k,2}$ – notation is inconsistent...

Some people claim that W stand for Weyl, and H for Hardy or Hilbert.

Exercise 13.3. *For $s > 0$ let*

$$f(x) := |x|^{-s}.$$

Observe that f is only defined for $x \neq 0$, but since measurable functions need only be defined outside of a null-set this is still a reasonable function.

We have already seen, Exercise 3.50,that $f \in L^p_{loc}(\mathbb{R}^n)$ for any $1 \leq p < \frac{n}{s}$.

(1) Compute for $x \neq 0$ that

$$\partial_i f(x) = -s\,|x|^{-s-2} x^i \tag{13.1}$$

and show $\partial_i f \in L^q_{loc}(\mathbb{R}^n \setminus \{0\})$ for any $1 \leq q < \frac{n}{s+1}$.

(2) Show that (13.1) holds in the distributional sense, i.e. that if $n \geq 2$ and $0 < s < n-1$ then for any $\varphi \in C^\infty_c(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f(x)\,\partial_i\varphi(x)\,dx = \int_{\mathbb{R}^n} s\,|x|^{-s-2} x^i \varphi(x)\,dx.$$

(3) conclude that $f \in W^{1,q}_{loc}(\mathbb{R}^n)$ for any $1 \leq q < \frac{n}{s+1}$.

Exercise 13.4. *Let*

$$f(x) := \log|x|.$$

Show that $f \in L^p_{loc}(\mathbb{R}^n)$ for any $1 \leq p < \infty$, and $f \in W^{1,p}_{loc}(\mathbb{R}^n)$ for all $p \in [1, n)$, if $n \geq 2$.

Exercise 13.5. *Let*

$$f(x) := \log\log\frac{2}{|x|} \quad \text{in } B(0,1)$$

Show that for $n \geq 2$, $f \in W^{1,n}(B(0,1))$.

Moreover, for $n = 2$, in distributional sense

$$\Delta f = |Df|^2$$

Observe that this serves as an example for solutions to nice differential equations that are not continuous!

Exercise 13.6. *Show that $f(x) := \frac{x}{|x|}$ belongs to $W^{1,p}(\mathbb{B}^n, \mathbb{R}^n)$ whenever $p < n$.*

Proposition 13.7 (Basic properties of weak derivatives)**.** *Let $u, v \in W^{k,p}(\Omega)$ and $|\gamma| \leq k$. Then*

(1) $\partial^\gamma u \in W^{k-|\gamma|,p}(\Omega)$.
(2) Moreover $\partial^\alpha\partial^\beta u = \partial^\beta\partial^\alpha u = \partial^{\alpha+\beta} u$ if $|\alpha| + |\beta| \leq k$.

(3) *For each* $\lambda, \mu \in \mathbb{R}$ *we have* $\lambda u + \mu v \in W^{k,p}(\Omega)$ *and*
$$\partial^\alpha(\lambda u + \mu v) = \lambda \partial^\alpha u + \mu \partial^\alpha v$$
(4) *If* $\Omega' \subset \Omega$ *is open then* $u \in W^{k,p}(\Omega')$
(5) *For any* $\eta \in C_c^\infty(\Omega)$, $\eta u \in W^{k,p}$ *and (if* $k \geq 1$*), and we have the Leibniz formula (aka product rule)*
$$\partial_i(\eta u) = \partial_i \eta\ u + \eta \partial_i u.$$

Proof. (1) We show that $\partial_i u \in W^{k-1,p}(\Omega)$, only. The general statement then follows accordingly. By definition of the distributional derivative we have that $\partial_i u \in L^p(\Omega)$. For any $|\beta| \leq k-1$ and $\varphi \in C_c^\infty(\Omega)$ we have
$$\int_\Omega \partial_i u\, \partial^\beta \varphi = -\int_\Omega u\, \partial_i \partial^\beta \varphi = -(-1)^{|\beta|+1} \int_\Omega \partial_i \partial^\beta u\, \varphi = (-1)^{|\beta|} \int_\Omega \partial_i \partial^\beta u\, \varphi.$$
The first inequality comes from the fact that $\partial^\beta \varphi \in C_c^\infty(\Omega)$ and from the definition of the weak derivative ∂_i. The second equation comes from the definition of the weak derivative of $\partial_i \partial^\beta$ for $W^{k,p}$-functions.

(2) We show $\partial_i \partial_j u = \partial_j \partial_i u$, again the general case follows. And as above this is proven by deducing respective properties from the properties in the space of test-funtions: For $\varphi \in C_c^\infty(\Omega)$ we have $\partial_i \partial_j \varphi = \partial_j \partial_i \varphi$, and thus
$$\int_\Omega \partial_i \partial_j u\, \varphi = \int_\Omega u\, \partial_i \partial_j \varphi = \int_\Omega u\, \partial_j \partial_i \varphi = \int_\Omega \partial_j \partial_i u\, \varphi.$$
(3) Follows from the linearity of the definition of weak derivative and the equivalent statements for smooth functions $\varphi \in C_c^\infty(\Omega)$

(4) If $\Omega' \subset \Omega$ then any $\varphi \in C_c^\infty(\Omega')$ belongs also to $C_c^\infty(\Omega)$. That is any property true for test functions $\varphi \in C_c^\infty(\Omega)$ holds also for testfunctions in $\varphi \in C_c^\infty(\Omega')$.

(5) For $\varphi \in C_c^\infty(\Omega)$ we have by the usual Leibniz rule
$$\begin{aligned}\int_\Omega \eta u \partial_i \varphi &= \int_\Omega u\, \partial_i(\eta\varphi) - \int_\Omega u\, \partial_i \eta\ \varphi \\ &= -\int_\Omega \partial_i u\, \eta \varphi - \int_\Omega u\, \partial_i \eta\ \varphi \\ &= -\int_\Omega (\partial_i u\, \eta + u\, \partial_i \eta)\ \varphi\end{aligned}$$
The second equation is the definition of weak derivative $\partial_i u$ (since $\eta\varphi \in C_c^\infty(\Omega)$ is a permissible testfunction).

That is we have shown for all $\varphi \in C_c^\infty(\Omega)$,
$$\int_\Omega \eta u \partial_i \varphi = \int_\Omega u\, \partial_i(\eta\varphi) - \int_\Omega u\, \partial_i \eta\ \varphi.$$
This means that in distributional sense $\partial_i(\eta u) = \partial_i \eta\ u + \eta \partial_i u$. Now observe that $\eta u \in L^p(\Omega)$ and $\partial_i \eta\ u + \eta \partial_i u \in L^p(\Omega)$, so $\eta u \in W^{1,p}(\Omega)$.

□

Proposition 13.8. *$(W^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$, $(H^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$, $(H_0^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$ are all Banach spaces.*

For $p = 2$ they are Hilbert spaces, with inner product

$$\langle u, v\rangle = \sum_{|\gamma|\leq k} \int \partial^\gamma u\, \partial^\gamma v.$$

Proof. $\|\cdot\|_{W^{k,p}(\Omega)}$ is a norm. By definition $(H^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$, $(H_0^{k,p}(\Omega), \|\cdot\|_{W^{k,p}(\Omega)})$ are complete and thus Banach spaces.

As for the completeness of $W^{k,p}(\Omega)$, it essentially follows from the completeness of $L^p(\Omega)$.

Let $(u_i)_{i\in\mathbb{N}} \subset W^{k,p}(\Omega)$ be a Cauchy sequence of $W^{k,p}$-functions, i.e.

$$\forall \varepsilon > 0 \exists N = N(\varepsilon) \in \mathbb{N} \quad \text{s.t. } \forall i, j \geq N: \quad \|u_i - u_j\|_{W^{k,p}(\Omega)} < \varepsilon.$$

We have to show that u_i converges to some $u \in W^{k,p}(\Omega)$ in the $W^{k,p}(\Omega)$-norm.

Observe that by the definition of the $W^{k,p}$-norm, if u_i is a Cauchy sequence for $W^{k,p}$, then for any $|\gamma| \leq k$, $(\partial^\gamma u_i)_{i\in\mathbb{N}}$ are Cauchy sequences of $L^p(\Omega)$.

Since $L^p(\Omega)$ is a Banach space, i.e. complete, each $\partial^\gamma u_i$ converges in $L^p(\Omega)$ to some object which we call $\partial^\gamma u$,

$$\|\partial^\gamma u_i - \partial^\gamma u\|_{L^p(\Omega)} \xrightarrow{i\to\infty} 0 \quad \forall |\gamma| \leq k.$$

Observe that as of now we do not know that $\partial^\gamma u$ is actually the weak derivative of u! But we can check this is the case.

Since $\partial^\gamma u_i$ is the weak derivative of u_i, we have

$$\int_\Omega \partial^\gamma u_i\, \varphi = (-1)^{|\gamma|} \int_\Omega u_i\, \partial^\gamma \varphi \quad \forall \varphi \in C_c^\infty(\Omega).$$

But on both sides we have strong convergence in $L^p(\Omega)$. For any (fixed) $\varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega \partial^\gamma u\, \varphi \xleftarrow{i\to\infty} \int_\Omega \partial^\gamma u_i\, \varphi = (-1)^{|\gamma|} \int_\Omega u_i\, \partial^\gamma \varphi \xrightarrow{i\to\infty} (-1)^{|\gamma|} \int_\Omega u\, \partial^\gamma \varphi$$

and thus for any $\varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega \partial^\gamma u\, \varphi = (-1)^{|\gamma|} \int_\Omega u\, \partial^\gamma \varphi.$$

That is, $\partial^\gamma u$ is indeed the weak derivative of u, thus $u \in W^{k,p}(\Omega)$ and by the definition of the $W^{k,p}$-norm

$$\|u_i - u\|_{W^{k,p}(\Omega)} \xrightarrow{i\to\infty} 0.$$

$\square$

13.1. **Approximation by smooth functions.** We mentioned above the $H = W$ problem, i.e. we would like to approximate Sobolev functions by smooth functions. Why? Because then we don't have to deal that many times with the weak definition of derivatives, but show desired results for smooth functions, then pass to the limit and hopefully obtain the result for Sobolev maps. Observe that since $W^{k,p}(\Omega)$ is a Banach space, and $C^\infty(\overline{\Omega}) \subset W^{k,p}(\Omega)$ (exercise!) we clearly have $H^{k,p}(\Omega) \subset W^{k,p}(\Omega)$. for the other direction we now obtain the first result:

Proposition 13.9 (Local approximation by smooth functions). *Let $u \in W^{k,p}(\Omega)$, $1 \leq p < \infty$. Set*

$$u_\varepsilon(x) := \eta_\varepsilon * u(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(y-x)\, u(y)\, dy.$$

Here $\eta_\varepsilon(z) = \varepsilon^{-n}\eta(z/\varepsilon)$ for the usual bump function $\eta \in C_c^\infty(B(0,1),[0,1])$, $\int_{B(0,1)} \eta = 1$. Then

(1) $u_\varepsilon \in C^\infty(\Omega_{-\varepsilon})$, where as before

$$\Omega_{-\varepsilon} := \{x \in \Omega : \operatorname{dist}(x, \partial\Omega) > \varepsilon\}$$

for each $\varepsilon > 0$ such that $\Omega_{-\varepsilon} \neq \emptyset$.

(2) Moreoever for any $\Omega' \subset\subset \Omega$,

$$\|u_\varepsilon - u\|_{W^{k,p}(\Omega')} \xrightarrow{\varepsilon \to 0} 0.$$

Proof. (1) As in Theorem 4.25 we have $u_\varepsilon \in C^\infty(\Omega_{-\varepsilon})$ – we do not need that u is a Sobolev function, but merely that $u \in L^p(\Omega)$.

(2) Next we claim that $\partial^\gamma u_\varepsilon(x) = (\partial^\gamma u)_\varepsilon(x)$ for $x \in \Omega_-$. Indeed, for $x \in \Omega_{-\varepsilon}$,

$$\partial^\gamma u_\varepsilon(x) = \int_\Omega \partial_x^\gamma \left(\eta_\varepsilon(x-z)\right) u(z)\, dz = (-1)^{|\gamma|} \int_\Omega \partial_z^\gamma \left(\eta_\varepsilon(x-z)\right) u(z)\, dz.$$

Now we observe that $\eta_\varepsilon(x - \cdot) \in C_c^\infty(\Omega)$ if $x \in \Omega_{-\varepsilon}$: observing size of the support of η_ε, $\operatorname{supp} \eta_\varepsilon \subset B(0,\varepsilon)$.

Thus by the definition of weak derivative,

$$(-1)^{|\gamma|} \int_\Omega \partial_z^\gamma \left(\eta_\varepsilon(x-z)\right) u(z)\, dz = \int_\Omega \eta_\varepsilon(x-z) \partial^\gamma u(z)\, dz = (\partial^\gamma u)_\varepsilon(x).$$

Now, for any $\Omega' \subset\subset \Omega$ and $\varepsilon < \operatorname{dist}(\Omega', \partial\Omega)$, for any $1 \leq p < \infty$[30]

$$\|(\partial^\gamma u)_\varepsilon - \partial^\gamma u\|_{L^p(\Omega')} \xrightarrow{\varepsilon \to 0} 0.$$

This holds for any γ such that $\partial^\gamma u \in L^p$, i.e. for all $|\gamma| \leq k$. We conclude that

$$\|u_\varepsilon - u\|_{W^{k,p}(\Omega')} \xrightarrow{\varepsilon \to 0} 0.$$

□

[30] but not for $p = \infty$!

Even though Proposition 13.9 is only about local approximation, it is very useful to prove properties of Sobolev function.

Lemma 13.10. *For $1 \leq p < \infty$[31], if $v \in W^{1,p}(\Omega)$ and $f \in C^1(\mathbb{R},\mathbb{R})$ with $[f]_{\mathrm{Lip}\,(\mathbb{R})} \equiv \|f'\|_{L^\infty(\mathbb{R}^n)} < \infty$ then $f(v) \in W^{1,p}(\Omega)$, and we have in distributional sense*

$$\partial_\alpha(f(v)) = f'(v)\,\partial_\alpha v. \tag{13.2}$$

Proof. Let v_ε be the (local) approximation of v in $W^{1,p}_{loc}(\Omega)$ from Proposition 13.9.

First we observe that (13.2) is true if v was a differentiable function, in particular,

$$\partial_\alpha(f(v_\varepsilon)) = f'(v_\varepsilon)\,\partial_\alpha v_\varepsilon \quad \text{in } \Omega_{-\varepsilon}.$$

Now let $\varphi \in C_c^\infty(\Omega)$, and take ε_0 so small such that $\Omega' := \operatorname{supp}\varphi \subset \Omega_{-\varepsilon}$ for all $\varepsilon \in (0,\varepsilon_0)$. Then we have for all $\varepsilon < \varepsilon_0$,

$$\int_\Omega f(v_\varepsilon)\,\partial_\alpha\varphi = -\int_\Omega f'(v_\varepsilon)\,\partial_\alpha v_\varepsilon\,\varphi \tag{13.3}$$

Now we observe that $f(v_\varepsilon) \xrightarrow{\varepsilon\to 0} f(v)$ with respect to the $L^p(\Omega')$-norm. Indeed, observe that $\Omega' \subset\subset \Omega$, so by Proposition 13.9,

$$\|f(v_\varepsilon) - f(v)\|_{L^p(\Omega')} \leq \|f'\|_{L^\infty(\Omega')}\,\|v_\varepsilon - v\|_{L^p(\Omega')} \xrightarrow{\varepsilon\to 0} 0.$$

That is, the left-hand side of (13.3) converges (recall $\operatorname{supp}\partial_\alpha\varphi \subset \operatorname{supp}\varphi = \Omega'$)

$$\int_\Omega f(v)\,\partial_\alpha\varphi \equiv \int_{\Omega'} f(v)\,\partial_\alpha\varphi = \lim_{\varepsilon\to 0}\int_{\Omega'} f(v_\varepsilon)\,\partial_\alpha\varphi \equiv \lim_{\varepsilon\to 0}\int_\Omega f(v_\varepsilon)\,\partial_\alpha\varphi.$$

As for the right-hand side of (13.3) we have that $\partial_\alpha v_\varepsilon \xrightarrow{\varepsilon\to 0} v$ in $L^p(\Omega')$, and $f'(v_\varepsilon) \xrightarrow{\varepsilon\to 0} f'(v)$ almost everywhere in Ω (up to taking a subsequence $\varepsilon \to 0$)[32]. By dominated convergence, Theorem 3.26, this implies

$$\int_\Omega f'(v_\varepsilon)\,\partial_\alpha v_\varepsilon\,\varphi \xrightarrow{\varepsilon\to 0} \int_\Omega f'(v)\,\partial_\alpha v\,\varphi$$

Then from (13.3) we get the claim, observing that $f'(v)\partial_\alpha v \in L^p(\Omega)$, since $f' \in L^\infty$. □

Remark 13.11. Actually, a stronger statement is true: if $u \in W^{1,p}(\Omega)$ and $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, $f \in C^{0,1}$ then $f \circ u \in W^{1,p}(\Omega)$. Again, formally this is looks easy since $\nabla(f \circ u) = Df(u)\,\nabla u$– since f is almost everywhere differentiable.

We first just sketch the proof of a special case:

[31] we can later conclude, using Theorem 13.24, that this also holds for $p = \infty$, since then Sobolev maps are simply Lipschitz maps

[32] these are results from measure theory: since f' is continuous, and since L^1-convergence implies almost everywhere convergence up to subsequence, Theorem 3.51

Lemma 13.12. *Let $u \in W^{1,1}(\Omega)$, then $|u| \in W^{1,1}(\Omega)$.*

Moreover we have $Du = 0$ almost everywhere in $\{u(x) = 0\}$[33].

Also we have

$$D|u| = \frac{u}{|u|} Du.$$

Proof. We only sketch the proof.

The difficulty lies in the fact that $|\cdot|$ is merely Lipschitz continuous, so we mollify it:

$$f_{\varepsilon,\theta}(t) := \sqrt{(t+\theta\varepsilon)^2 + \varepsilon^2} - \sqrt{(\theta\varepsilon)^2 + \varepsilon^2}.$$

$f_{\varepsilon,\theta}$ is a smooth function.

One approximates $|u|$ by $u_\varepsilon := f_{\varepsilon,\theta}(u)$ for some $\theta \in \mathbb{R}$

Since $f_{\varepsilon,\theta}$ is smooth we have in distributional sense, by Lemma 13.10,

$$Du_\varepsilon = \frac{u + \varepsilon\theta}{\sqrt{(u+\varepsilon\theta)^2 + \varepsilon^2}} Du \xrightarrow{\varepsilon \to 0} Du \cdot \begin{cases} 1 & \text{in } \{u > 0\} \\ \frac{\theta}{\sqrt{\theta^2+1}} & \text{in } \{u = 0\} \\ -1 & \text{in } \{u < 0\} \end{cases}$$

Now $u_\varepsilon \to |u|$ in $L^1(\Omega)$, and Du_ε converges also in $L^1(\Omega)$. Using test functions and the convergence as $\varepsilon \to 0$ we get that

$$L^1(\Omega) \ni D|u| = Du \cdot \begin{cases} 1 & \text{in } \{u > 1\} \\ \frac{\theta}{\sqrt{\theta^2+1}} & \text{in } \{u = 0\} \\ -1 & \text{in } \{u < 1\} \end{cases}$$

But weak derivatives are unique as L^1-functions. The nonunique looks independent in θ. This means either $Du = 0$ almost everywhere in $\{u = 0\}$ or $\{u = 0\}$ is a zeroset (which still means that $Du = 0$ almost everywhere in $\{u = 0\}$). $\square$

Exercise 13.13. *Show that Lemma 13.12 does not hold in the other direction, i.e. there exist functions $u \in L^1(\mathbb{R})$ such that $|u| \in W^{1,1}(\mathbb{R})$ but $u \notin W^{1,1}(\mathbb{R})$.*

Hint: Example 4.47, see also Exercise 13.19

In general, crazy sets, it might be difficult to extend Proposition 13.9 to the boundary (think of an open set whose boundary is the Koch-curve, or an open set whose boundary has positive $\mathcal{L}^n$-measure!). To rule this out we make the following definition of C^k-boundary data

[33]Check this for smooth functions: Either $\{u(x) = 0\}$ is a zeroset. On the other hand, on the "substantial" parts of $\{u(x) = 0\}$ we should think of u as constant

Definition 13.14 (Regularity of boundary of sets)**.** Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that $\partial\Omega \in C^k$ (more generally in $C^{k,\alpha}$) if $\partial\Omega \subset \mathbb{R}^n$ is a C^k (or $C^{k,\alpha}$, respectively) manifold, that is if

for any $x \in \partial\Omega$ there exists a radius $r > 0$ and a C^k-diffeomorphism $\Phi : B(x,r) \to B(0,r)$ (i.e. the map Φ is a bijection between $B(x,r)$ and $B(0,r)$ and Φ and Φ^{-1} are both of class C^k) such that

- $\Phi(x) = 0$
- $\Phi(\Omega \cap B(x,r)) = B(0,r) \cap \mathbb{R}^n_+$
- $\Phi(B(x,r)\backslash\Omega) = B(0,r) \cap \mathbb{R}^n_-$.

Theorem 13.15 (Smooth approximation for Sobolev functions)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, and $\partial\Omega \in C^1$. For any $u \in W^{k,p}(\Omega)$ there exist a smooth approximating sequence $u_i \in C^\infty(\overline{\Omega})$ such that*
$$\|u_i - u\|_{W^{k,p}(\Omega)} \xrightarrow{i\to\infty} 0.$$

Proof. First we consider the situation close to the boundary.

Let $x_0 \in \partial\Omega$.

Observe first the following: If $x \in B(0,r)^+$ and $|z - x| < \varepsilon$ (for $\varepsilon \ll r$) then $x + \varepsilon e_n \subset B(x,2r)^+$. Since $\partial\Omega$ belongs to C^1 one can show that the same holds (on sufficiently small balls $B(x_0,r)$) as well: For some $\lambda = \lambda(x_0)$, a unit vector $\nu = \nu(x_0)$, if $z \in B(x,\varepsilon)$ and $x \in \Omega \cap B(x_0,r)$ then
$$z + \lambda\varepsilon\nu \in \Omega.$$
One should think of ν the inwards facing unit normal at x_0 (which can be computed from the derivatives of Φ and is continuous around x_0).

That is for $x \in \Omega' := B(x_0,r/2) \cap \Omega$ we may set
$$u_\varepsilon(x) := \int_{\mathbb{R}^n} \eta_\varepsilon(z-x)\, u(z+\lambda\nu\varepsilon)dz = \int_{\mathbb{R}^n} \eta_\varepsilon(z-\lambda\nu\varepsilon-x)\, u(z)dz.$$
Clearly, u_ε is still smooth, but now in all of of $\overline{\Omega'}$. Moreover observe that if we set
$$v_\varepsilon(x) := u(z+\lambda\nu\varepsilon).$$
we have
$$\|v_\varepsilon - u\|_{W^{k,p}(\Omega')} \xrightarrow{\varepsilon\to 0} 0$$
since v_ε is merely a translation. Moreover, $u_\varepsilon = \eta_\varepsilon * v_\varepsilon$, and thus as before
$$\|u_\varepsilon - v_\varepsilon\|_{W^{k,p}(\Omega')} =\xrightarrow{\varepsilon\to 0} 0.$$
We conclude that $u_\varepsilon \to u$ in $W^{k,p}(\Omega')$.

Now we cover all of $\partial\Omega$ by (finitely many, by compactness) balls $B(x_i,r_i)$ and choose the approximation $u_{\varepsilon,i}$ on $\Omega_i := B(x_i,r_i) \cap \Omega$ as above. In $\Omega_0 := \Omega\backslash \bigcup B(x_i,r_i) \subset\subset \Omega$ we can find another approximation $u_{\varepsilon,0}$.

Now we pick a smooth decomposition of unity η_i with support in $\Omega_i \cap \partial\Omega$ such that

$$\sum_{i\in\mathbb{N}} \eta_i \equiv 1 \quad \text{in } \Omega.$$

Setting

$$u_\varepsilon := \sum_i \eta_i u_{\varepsilon,i} \in C^\infty(\overline{\Omega}).$$

We then use the Leibniz rule to conclude that

$$\|u_\varepsilon - u\|_{W^{k,p}(\Omega)} \xrightarrow{\varepsilon\to 0} 0.$$

□

Theorem 13.15 can be improved to Lipschitz domains, cf. Exercise 13.29, but not necessarily to more irregular sets, Exercise 13.29. Observe that still $C^\infty(\Omega)$ (not $C^\infty(\overline{\Omega})$) is always dense in $W^{k,p}(\Omega)$:

Theorem 13.16. *Let Ω be open. Then $C^\infty(\Omega)$ (not necessarily $C^\infty(\overline{\Omega})$) is dense in $W^{k,p}(\Omega)$. By this we mean that, for any $f \in W^{k,p}(\Omega)$ and any $\varepsilon > 0$ there exists $f_\varepsilon \in C^\infty(\Omega) \cap W^{k,p}(\Omega)$ such that*

$$\|f_\varepsilon - f\|_{W^{k,p}(\Omega)} < \varepsilon.$$

Proof. We are not given the full proof here (for this see, e.g., [Adams and Fournier, 2003, Theorem 3.17]) but only the idea:

We know that any open set can be written as countable union of closed dyadic cubes,

$$\Omega = \bigcup_{i=1}^\infty Q_i.$$

where the cubes' interior is pairwise disjoint. We can refine these cubes into what is called *Whitney cubes* or *Whitney decomposition* of Ω, [Grafakos, 2014, Appendix J.1]: far away from the boundary we take large cube, towards the boundary we take smaller cubes. This way we can ensure the folllowing property:

- for any cube we have

$$\sqrt{n}\text{sidelength}(Q_j) \le \text{dist}\,(Q_j, \mathbb{R}^n \setminus \Omega).$$

- whenever two cubes Q_j and Q_k touch (i.e. their boundary) then

$$\frac{1}{4} \le \frac{\text{sidelength}(Q_j)}{\text{sidelength}(Q_k)} \le 4.$$

- Each Q_j touches at most $12^n - 4^n$ other cubes Q_k

In particular, if we denote by $Q_j^* \subset Q_j^{**}$ are cubes with the same center as Q_j but slightly increased sidelength (e.g. sidelength$(Q_j^*) = \frac{17}{16}$sidelength(Q_j) and sidelength$(Q_j^{**}) = \frac{18}{16}$sidelength$(Q_j)$) then each Q_j^*, Q_j^{**} is still contained in Ω. Moreover any Q_j^* it intersects with at most 12^n-4^n other cubes Q_k^*, and likewise any Q_j^{**} it intersects with at most $12^n - 4^n$ other cubes Q_k^{**}

Now we can find a decomposition of unity $\eta_j \in C_c^\infty(Q_j^*)$ such that $\eta_j \equiv 1$ in Q_j, and for any $x \in \Omega$,

$$1 = \sum_j \eta_j(x) \quad \text{and the sum is finite.}$$

Let now $f \in W^{k,p}(\Omega)$ and fix $\varepsilon > 0$.

Then $\eta_j f \in W^{k,p}(\Omega)$, with $\operatorname{supp} \eta_j f \subset Q_j^* \subset\subset \Omega$. As in Proposition 13.9 we can then find $g_j \in C_c^\infty(Q_j^{**})$ such that

$$\|g_j - f\|_{W^{k,p}(\Omega)} = \|g_j - f\|_{W^{k,p}(Q_j^*)} < 2^{-j}\varepsilon.$$

Now let

$$f_\varepsilon(x) := \sum_j g_j(x)$$

Observe that since each g_j is supported in Q_j^{**} (which only intersects with at most $12^n - 4^n$ many other Q_k^{**}) this sum is locally finite – and thus $f_\varepsilon \in C^\infty(\Omega)$ (but not necessarily $f_\varepsilon \in C^\infty(\overline{\Omega})$!). Then

$$\|f_\varepsilon - f\|_{L^p(\Omega)} \leq \sum_j \|g_j - f\|_{L^p(Q_j^*)} \leq \sum_j 2^{-j}\varepsilon = \varepsilon.$$

and similar for all derivatives. □

Observe that Theorem 13.16 can be used to show completeness of $W^{k,p}(\Omega)$ for any open set Ω, since we can write $W^{k,p}(\Omega)$ as the closure of smooth maps.

Exercise 13.17. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open sets, and let $\phi : \Omega' \to \Omega$ be a C^∞-diffeomorphism. Let $f : \Omega \to \mathbb{R}$ be measurable. Show that for all $k = \{0, 1, \ldots, \}$ and all $p \in [1, \infty]$*

$$f \in W^{k,p}(\Omega) \quad \Leftrightarrow f \circ \Phi \in W^{k,p}(\Omega').$$

On $\mathbb{R}^n$ approximation is much easier, indeed we can approximate with respect to the $W^{k,p}$-norm any $u \in W^{k,p}(\mathbb{R}^n)$ by functions $u_k \in C_c^\infty(\mathbb{R}^n)$. That is, $W^{k,p}(\mathbb{R}^n) = W_0^{k,p}(\mathbb{R}^n)$. We could describe this as "$u \in W^{k,p}(\mathbb{R}^n)$ implies that u and $k-1$-derivatives of u all vanish at infinity".

Proposition 13.18. *(1) Let $u \in W^{k,p}(\Omega)$, $p \in [1, \infty)$. If $\operatorname{supp} u \subset\subset \Omega$ then there exists $u_k \in C_c^\infty(\Omega)$ such that*

$$\|u - u_k\|_{W^{k,p}(\Omega)} \xrightarrow{k\to\infty} 0.$$

(2) Let $u \in W^{k,p}(\mathbb{R}^n)$, $p \in [1, \infty)$. Then there exists $u_k \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|u - u_k\|_{W^{k,p}(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

(3) Let $u \in W^{k,p}(\mathbb{R}^n_+) = \mathbb{R}^{n-1} \times (0,\infty))$. Then there exists $u \in C_c^\infty(\mathbb{R}^{n-1} \times [0,\infty)$ (i.e., u may not be zero on $(x',0)$ for small x') such that

$$\|u - u_k\|_{W^{k,p}(\mathbb{R}^n_+)} \xrightarrow{k\to\infty} 0.$$

Proof. (1) follows from the proof of Proposition 13.9: Observe that $\operatorname{supp} u \subset\subset \Omega$ implies that $\eta_\varepsilon * u \in C_c^\infty(\Omega)$ if ε is only small enough.

(3) is an exercise, a combination of the proof of (2) and Theorem 13.15.

So let us discuss (2). Let $\eta \in C_c^\infty(B(0,1))$ again be the typical mollifier bump function, $\eta_\varepsilon(x) = \varepsilon^{-n}\eta(x/\varepsilon)$. We have already seen that

$$\eta_\varepsilon * u \in C^\infty(\mathbb{R}^n).$$

But there is no reason that $\eta_\varepsilon * u \in C_c^\infty(\mathbb{R}^n)$. Set (without rescaling by R^n!)

$$\varphi_R(x) := \eta(x/R) \in C_c^\infty(B(0,R)).$$

Then we set

$$u_{\varepsilon,R} := \eta_\varepsilon * (\varphi_R\, u)$$

Now before $u_{\varepsilon,R} \in C_c^\infty(B(0,R+\varepsilon)) \subset C_c^\infty(\mathbb{R}^n)$.

Moreover we have for any $\ell = 0,\ldots,k$

$$\begin{aligned}
\|\nabla^\ell(u - u\varphi_R)\|_{L^p(\mathbb{R}^n)} =&\|\nabla^\ell(1-\varphi_R)u\|_{L^p(\mathbb{R}^n)}\\
\le& C(\ell)\sum_{i=0}^{\ell}\|\nabla^i(1-\varphi_R)\nabla^{\ell-i}u\|_{L^p(\mathbb{R}^n)}\\
\le& C(\ell)\,\|(1-\varphi_R)\nabla^\ell u\|_{L^p(\mathbb{R}^n)} + C(\ell)\sum_{i=0}^{\ell}\|\nabla^i(1-\varphi_R)\|_\infty\|\nabla^{\ell-i}u\|_{L^p(\mathbb{R}^n)}\\
\le& C(\ell,\eta)\,\|(1-\varphi_R)\nabla^\ell u\|_{L^p(\mathbb{R}^n\setminus B(0,R))} + C(\ell,\eta)\sum_{i=0}^{\ell}R^{-i}\|u\|_{W^{k,p}(\mathbb{R}^n)}\\
\xrightarrow{R\to\infty}&0
\end{aligned}$$

by Lebesgue dominated convergence theorem.

On the other hand, as already seen, for $R > 0$ fixed,

$$\|\eta_\varepsilon * (\varphi_R\, u) - \varphi_R u\|_{W^{k,p}(\mathbb{R}^n)} \xrightarrow{\varepsilon\to 0} 0.$$

Now we show that for any $\ell > 0$ there exists ε_ℓ, R_ℓ such that for $u_\ell := u_{\varepsilon_\ell,R_\ell}$ we have

$$\|u_\ell - u\|_{W^{k,p}(\mathbb{R}^n)} < \frac{1}{\ell} \xrightarrow{\ell\to\infty} 0, \tag{13.4}$$

that is $C_c^\infty(\mathbb{R}^n) \ni u_\ell \to u$ in $W^{k,p}(\mathbb{R}^n)$.

First, by the arguments above we can choose R_ℓ large enough such that

$$\|u - u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)} \le \frac{1}{2\ell}.$$

Next, we can choose ε_ℓ small enough such that

$$\|\eta_{\varepsilon_\ell} * (u\varphi_{R_\ell}) - u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)} \le \frac{1}{2\ell}.$$

Thus, by triangular inequality,

$$\|u_\ell - u\|_{W^{k,p}(\mathbb{R}^n)} \le \underbrace{\|u_\ell - u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)}}_{=\|\eta_{\varepsilon_\ell}*(u\varphi_{R_\ell})-u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)}} + \|u\varphi_{R_\ell} - u\|_{W^{k,p}(\mathbb{R}^n)} \le 2\frac{1}{2\ell} = \frac{1}{\ell}.$$

This proves (13.4), and thus (2) is established. □

Exercise 13.19. *Let*

$$f(x) := \begin{cases} 1 & x_1 \ge 0 \\ 0 & x_1 < 0 \end{cases}$$

Use a mollification argument to show that f does not belong to $W^{1,p}_{loc}(\mathbb{R}^n)$.

See also Example 4.47 for another argument.

Exercise 13.20 (Censored Mollification)**.** *Take three radii $0 < r < \rho < R$ and assume $u \in W^{\ell,p}(B(0,R))$, $p \in [1,\infty)$, $\ell \in \mathbb{N} \cup \{0\}$. Show that there is an approximation $u_k \in W^{\ell,p}(B(0,R))$ with*

$$\|u_k - u\|_{W^{\ell,p}(B(0,R)} \xrightarrow{k\to\infty} 0$$

that satisfies the following conditions for all $k \in \mathbb{N}$

- $u_k \equiv u$ *in* $B(0,R) \setminus B(0,\rho)$
- $u_k \in C^\infty(B(0,r))$

Hint: Use the same approximation as in Exercise 4.39

Let $\Omega, \Omega_1, \Omega_2$ be three open, connected, and bounded sets in $\mathbb{R}^n$ such that $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$ and $\Omega_1 \cap \Omega_2 = \emptyset$. Let $u \in L^p(\Omega)$ such that $u\Big|_{\Omega_1} \in W^{1,p}(\Omega_1)$ and $u\Big|_{\Omega_2} \in W^{1,p}(\Omega_2)$. Then there is – in general – no reason that $u \in W^{1,p}(\Omega)$. Indeed, take $u \equiv 1$ in Ω_1 and $u \equiv 0$ in Ω_2, which in general will not belong to $W^{1,p}(\Omega)$ because of the jump. It is a nice fact (used in numerics) that the jump is indeed all that can go wrong, in the following sense.

Lemma 13.21. *Let $\Omega, \Omega_1, \Omega_2$ be open, connected, and bounded sets as above, all with Lipschitz boundary. Assume $u \in L^p(\Omega)$ such that $u\Big|_{\Omega_1} \in W^{1,p}(\Omega_1)$ and $u\Big|_{\Omega_1} \in W^{1,p}(\Omega_2)$.*

If u is moreover continuous in Ω, then $u \in W^{1,p}(\Omega)$.

Proof. We only prove this in the situation where $\Omega = \mathbb{R}^n$, $\Omega_1 = \mathbb{R}^{n-1} \times (0,\infty)$, $\Omega_2 = \mathbb{R}^{n-1} \times (-\infty, 0)$. Observe that we can set

$$Du(x) := \begin{cases} Du(x) & x \in \Omega_1 \\ Du(x) & x \in \Omega_2 \end{cases}$$

Since $\partial\Omega_1, \partial\Omega_2$ are zerosets, this means that $Du \in L^p(\Omega)$. *However* this does not mean that $Du(x)$ is actually the (distributional!) derivative of u. This we still have to show (and this is where continuity of u plays a role).

Formally the argument goes as follows: Let $\nu = (0, 0, \ldots, 0, 1)$. Integration by parts,

$$\begin{aligned}\int_{\mathbb{R}^n} u\partial_\alpha\varphi - \int_{\mathbb{R}^n} \partial_\alpha u\varphi &= \int_{\mathbb{R}^{n-1}\times\{0\}} u(-\nu^\alpha)\varphi + \int_{\mathbb{R}^{n-1}\times\{0\}} u(\nu^\alpha)\varphi \\ &= \int_{\mathbb{R}^{n-1}\times\{0\}} (u-u)(\nu^\alpha)\varphi = 0.\end{aligned}$$

What have we used? We have used that

$$\int_{\mathbb{R}^n_\pm} u\partial_\alpha\varphi - \int_{\mathbb{R}^n_\pm} \partial_\alpha u\varphi = \int_{\mathbb{R}^n_+} u(\mp\nu^\alpha)\varphi.$$

This can be proven (by approximation) since u is indeed continuous at the boundary $\mathbb{R}^{n-1} \times \{0\} = \partial\mathbb{R}^n_+ = \partial\mathbb{R}^n_-$ (more precisely we need that u is really defined at the boundary. □

13.2. **Difference Quotients.** In PDE one likes to use the method of differentiating the equation (e.g. that if $\Delta u = 0$ then also for $v := \partial_i u$ we have $\Delta v = 0$ – so we can easier estimates for $\partial_i u$). In the Sobolev space category this is also a useful technique. Sometimes, the "first assume that everything is smooth, then use mollification"-type argument is difficult to put into practice. In this case, a technique developed by Nirenberg, is discretely differentiating the equation (which does not require the function to be a priori differentiable):

$$\Delta u = 0 \Rightarrow v(x) := (\Delta_h^{e_i} u)(x) := \frac{u(x+he_i) - u(x)}{h} : \quad \Delta v = 0$$

For this to work, we need some good estimates. Recall that (by the fundamental theorem of calculus), for C^1-functions u,

$$\|\Delta_h^{e_\ell} u\|_{L^\infty} \leq \|\partial_\ell u\|_{L^\infty}.$$

This also holds in L^p for $W^{1,p}$-functions u, which is a result attributed to Nirenberg, see Proposition 13.23.

One important ingredient is that the fundamental theorem of calculus holds for Sobolev functions:

Lemma 13.22. *Let $u \in W^{1,1}_{loc}(\Omega)$. Fix $v \in \mathbb{R}^n$. Then for almost every $x \in \Omega$ such that the path $[x, x+v] \subset \Omega$ we have*

$$u(x+v) - u(x) = \int_0^1 \partial_\alpha u(x+tv) v^\alpha \, dt.$$

Proof. Let $\Omega' \subset \Omega$. In view of Proposition 13.9 we can approximate u by $u_k \in C^\infty(\overline{\Omega'})$ such that

$$\|u_k - u\|_{W^{1,1}(\Omega')} \xrightarrow{k\to\infty} 0.$$

The claim holds for the smooth functions u_k, namely we have that whenever $[x, x+v] \subset \Omega'$,

$$u_k(x+v) - u_k(x) = \int_0^1 \partial_\alpha u_k(x+tv) v^\alpha \, dt. \tag{13.5}$$

Now we have

$$\|u_k(\cdot + v) - u_k(\cdot) - (u(\cdot+v) - u(\cdot))\,\|_{L^p(\Omega')} \xrightarrow{k\to\infty} 0,$$

in particular (up to taking a subsequence),

$$u_k(x+v) - u_k(x) \xrightarrow{k\to\infty} u(x+v) - u(x) \quad \text{for almost every } x \in \Omega' \text{ such that } [x, x+v] \subset \Omega'.$$

Also the right-hand side converges. Observing that[34]

$$\left(\int_\Omega \left(\int_0^1 |f(x,t)|\, dt\right)^p\right)^{\frac{1}{p}} \le \left(\int_0^1 \int_\Omega |f(x,t)|^p \, dt\right)^{\frac{1}{p}}$$

we have

$$\begin{aligned} &\|\int_0^1 \partial_\alpha u_k(\cdot + tv) v^\alpha \, dt - \int_0^1 \partial_\alpha u(\cdot + tv) v^\alpha \, dt\|_{L^p(\Omega')} \\ \le &\left(\int_0^1 |v| \|(Du_k - Du\|^p_{L^p(\Omega')} \, dt\right)^{\frac{1}{p}} = \|Du_k - Du\|_{L^p(\Omega')} \xrightarrow{k\to\infty} 0. \end{aligned}$$

So again, up to possibly a subsubsequence,

$$\int_0^1 \partial_\alpha u_k(x+tv) v^\alpha \, dt \xrightarrow{k\to\infty} \int_0^1 \partial_\alpha u(x+tv) v^\alpha \, dt$$

for all x such that $[x, x+v] \subset \Omega$.

Taking the limit in (13.5) we conclude. □

Proposition 13.23. *(1) Let $k \in \mathbb{N}$, (i.e. $k \neq 0$), and $1 < p < \infty$. Assume that $\Omega' \subset\subset \Omega$ are two open (nonempty) sets, and let $0 < |h| < \operatorname{dist}(\Omega', \partial\Omega)$. For $u \in W^{k,p}(\Omega)$ we have*

$$\|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')} \le \|\partial_\ell u\|_{W^{k-1,p}(\Omega)}.$$

[34]This can be seen by Jensens inequalty: For any $p \in [1, \infty)$,

$$(\fint_A |f|)^p \le \fint_A |f|^p,$$

this can also be shown by Hölder's inequality. Then Fubini gets to the claim.

Moreover we have

$$\|\Delta_h^{e_\ell} u - \partial_\ell u\|_{W^{k-1,p}(\Omega')} \xrightarrow{h\to 0} 0.$$

(2) *Let $u \in W^{k-1,p}(\Omega)$, $1 < p \leq \infty$. Assume that for any $\Omega' \subset\subset \Omega$ and any $\ell = 1,\ldots,n$ there exists a constant $C(\Omega')$ such that*

$$\sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')} \leq C(\Omega',\ell)$$

Then we $u \in W^{k,p}_{loc}(\Omega)$, and for any $\Omega' \subset \Omega$ we have

$$\|\partial_\ell u\|_{W^{k-1,p}(\Omega')} \leq \sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')}. \tag{13.6}$$

If $p=\infty$ we even have $u \in W^{k,\infty}(\Omega)$ with the estimate

$$\|\partial_\ell u\|_{W^{k-1,\infty}(\Omega)} \leq \sup_{\Omega'\subset\subset\Omega}\ \sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,\infty}(\Omega')}. \tag{13.7}$$

Proof of Proposition 13.23(1). The proof of (1) is essentially the same as for differentiable function, we use the fundamental theorem of calculus.

By the fundamental theorem of calculus, Lemma 13.22,

$$\Delta_h^{e_\ell} u(x) = \frac{1}{h}\int_0^1 \frac{d}{dt}\left(u(x+the_\ell)\right)\,dt = \frac{1}{h}h\int_0^1 \partial_\ell u(x+the_\ell)\,dt = \int_0^1 \partial_\ell u(x+the_\ell)\,dt,$$

Similarly, for any $|\gamma| \leq k-1$, $\ell = 1,\ldots,n$

$$|\Delta_h^{e_\ell}\partial^\gamma u(x)| \leq \int_0^1 |\partial_\ell\partial^\gamma u(x+the_\ell)|.$$

Taking the L^p-norm, observing that[35]

$$\left(\int_\Omega \left(\int_0^1 |f(x,t)|\,dt\right)^p\right)^{\frac{1}{p}} \leq \left(\int_0^1\int_\Omega |f(x,t)|^p\,dt\right)^{\frac{1}{p}}$$

we have

$$\|\Delta_h^{e_\ell}\partial^\gamma u\|_{L^p(\Omega')} \leq \left(\int_0^1 \|\partial_\ell\partial^\gamma u(\cdot+the_\ell)\|^p_{L^p(\Omega')}\,dt\right)^{\frac{1}{p}}.$$

Now observe that by substitution and $|h| < \operatorname{dist}(\Omega',\partial\Omega)$,

$$\|\partial_\ell\partial^\gamma u(\cdot+the_\ell)\|_{L^p(\Omega')} = \|\partial_\ell\partial^\gamma u(\cdot)\|_{L^p(\Omega'+he_\ell)} \leq \|\partial_\ell\partial^\gamma u(\cdot)\|_{L^p(\Omega)}$$

Consequently, for any $|h| < \operatorname{dist}(\Omega',\partial\Omega)$.

$$\|\Delta_h^{e_\ell}\partial^\gamma u\|_{L^p(\Omega')} \leq \left(\int_0^1 \|u\|^p_{W^{k,p}(\Omega)}\,dt\right)^{\frac{1}{p}} = \|u\|_{W^{k,p}(\Omega)}.$$

[35]This can be seen by Jensens inequalty: For any $p \in [1,\infty)$,

$$(\fint_A |f|)^p \leq \fint_A |f|^p,$$

this can also be shown by Hölder's inequality. Then Fubini gets to the claim.

This shows the first part of (1). For the second part we observe that by the same fundamental theorem argument as above,

$$|\Delta_h^{e_\ell}\partial^\gamma u(x) - \partial_\ell\partial^\gamma u(x)| = \int_0^1 |\partial_\ell\partial^\gamma u(x+the_\ell) - \partial_\ell\partial^\gamma u(x)|dt.$$

As above we obtain

$$\||\Delta_h^{e_\ell}\partial^\gamma u - \partial_\ell\partial^\gamma u|\|_{L^p(\Omega')} \leq \left(\int_0^1 \|\partial_\ell\partial^\gamma u(\cdot+the_\ell) - \partial_\ell\partial^\gamma u\|^p_{L^p(\Omega')}dt\right)^{\frac{1}{p}}$$

We can conclude by Lebesgue dominated convergence theorem once we show that for all $t \in (0,1)$,

$$\|\partial_\ell\partial^\gamma u(\cdot+the_\ell) - \partial_\ell\partial^\gamma u\|_{L^p(\Omega')} \xrightarrow{h\to 0} 0.$$

To obtain this last fact, fix $\varepsilon > 0$, let Γ_ε be a smooth approximation of $\partial_\ell\partial^\gamma u$ with

$$\|\partial_\ell\partial^\gamma u - \Gamma_\varepsilon\|_{L^p(\Omega'')} < \varepsilon,$$

where $\Omega'' \subset\subset \Omega$ is a slighly larger set than Ω' (and we can assume that h is small so that $\Omega' + the_\ell \subset \Omega$. Then

$$\|\partial_\ell\partial^\gamma u(\cdot+the_\ell) - \partial_\ell\partial^\gamma u\|_{L^p(\Omega')} \leq 2\varepsilon + \|\Gamma_\varepsilon(\cdot+the_\ell) - \Gamma_\varepsilon\|_{L^p(\Omega')} \xrightarrow{h\to 0} 2\varepsilon.$$

Letting $\varepsilon \to 0$ we conclude. □

Proof of Proposition 13.23(2). First let us assume that $p < \infty$.

Assume that for all $\ell \in \{1,\ldots,n\}$ we have

$$\sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')} < \infty.$$

$$\Delta_{h_i}^{e_\ell}\partial^\gamma u \xrightarrow{i\to\infty} f_{\ell,\gamma} \quad \text{weakly in } L^p(\Omega').$$

Since we are optimists, we call $\partial_\ell\partial^\gamma u := f_{\ell,\gamma} \in L^p(\Omega')$. We still need to show that $\partial_\ell\partial^\gamma u$ is actually the distributional derivative of u! Also, for simplicity of notation we drop the i in h_i and write $h \to 0$ (meaning always this subsequence). Weak convergence means in particular, that for any $\varphi \in C_c^\infty(\Omega') \subset L^p(\Omega')$,

$$\int_{\Omega'} \Delta_h^{e_\ell}\partial^\gamma u\ \varphi \xrightarrow{h\to 0} \int_{\Omega'} \partial_\ell\partial^\gamma u\ \varphi \tag{13.8}$$

Since $\operatorname{supp}\varphi \subset\subset \Omega'$ for $|h|$ small enough we have that $\Delta_{-h}^{e_\ell}\varphi \in C_c^\infty(\Omega')$. Now we perform a discrete integration by parts, namely by substitution,

$$\int_{\Omega'} \Delta_h^{e_\ell}\partial^\gamma u\ \varphi = -\int_{\Omega'} \partial^\gamma u\ \Delta_{-h}^{e_\ell}\varphi$$

Now since $\Delta_{-h}^{e_\ell}\varphi \in C_c^\infty(\Omega')$ is a testfunction and $u \in W^{k-1,p}$,

$$\int_{\Omega'} \Delta_h^{e_\ell}\partial^\gamma u\ \varphi = -\int_{\Omega'} \partial^\gamma u\ \Delta_{-h}^{e_\ell}\varphi = (-1)^{|\gamma|+1} - \int_{\Omega'} u\ \Delta_{-h}^{e_\ell}\partial^\gamma\varphi \xrightarrow{h\to 0} (-1)^{|\gamma|+1}\int_{\Omega'} u\ \partial_\ell\partial^\gamma\varphi,$$

in the last step we used dominated convergence and the smoothness of φ.

But then in (13.8) we obtain

$$(-1)^{|\gamma|+1}\int_{\Omega'} u\ \partial_\ell\partial^\gamma\varphi = \int_{\Omega'} \partial_\ell\partial^\gamma u\ \varphi$$

This holds for any $\ell \in \{1,\ldots,n\}$ and so we have shown that $\partial_\ell\partial^\gamma u$ is indeed the weak derivative of u which belongs to L^p, and thus $u \in W^{k,p}(\Omega')$. Since this holds for any $\Omega' \subset \Omega$ we conclude that $u \in W^{k,p}_{loc}(\Omega)$. The estimate (13.6) follows from the estimate of Theorem 12.13.

As for the case $p = \infty$, we observe first that for $\Omega' \subset\subset \Omega$ the estimate

$$\sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,\infty}(\Omega')} \leq C(\Omega',\ell) \tag{13.9}$$

implies (by Hölders inequality) also

$$\sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,2}(\Omega')} \leq C(\Omega',\ell)$$

Thus (13.9) implies $u \in W^{k,2}_{loc}(\Omega)$ and in view of Proposition 13.23(1) we have that $\Delta_h^{e_\ell} u \to \partial_\ell u$ in $W^{k-1,2}_{loc}(\Omega)$.

In particular, we already have the existence of the distributional derivative $\partial_\ell u \in W^{k-1,2}_{\text{loc}}(\Omega)$.

Set

$$\Lambda := \sup_{\Omega'\subset\subset\Omega}\ \sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,\infty}(\Omega')}.$$

For simplicity of notation in the following we shall assume $k = 1$.

We now claim that the above observations, together with (13.7), for any $\varphi \in C_c^\infty(\Omega)$,

$$\int_\Omega \partial_\ell u\,\varphi \leq \Lambda\,\|\varphi\|_{L^1(\Omega)}. \tag{13.10}$$

Indeed, since for $\varphi \in C_c^\infty(\Omega)$ let $\operatorname{supp}\varphi \subset \Omega' \subset\subset \Omega$, then we have

$$\int_\Omega \partial_\ell u\ \varphi = \lim_{|h|\to 0}\int_\Omega \underbrace{\Delta_h^{e_\ell} u}_{\leq\Lambda}\ \varphi \leq \Lambda\|\varphi\|_{L^1(\Omega)}.$$

which is exactly (13.10).

Let $x \in \Omega$ be a Lebesgue point of $\partial_\ell u$ in Ω, i.e.

$$\partial_\ell u(x) = \lim_{r\to 0} \fint_{B(x,r)} \partial_\ell u.$$

Observe that almost all points in Ω are Lebesgue points (since $\partial_\ell u \in L^2_{loc}(\Omega)$).

Set

$$\Omega' = \{z \in \Omega : \operatorname{dist}(x,\partial\Omega) < \frac{1}{2}\operatorname{dist}(x,\partial\Omega)\} \subset\subset \Omega.$$

Then for all $r < \frac{1}{4}\mathrm{dist}\,(x, \partial\Omega)$ we can set $\varphi := |B(x,r)|^{-1}\chi_{B(x,r)} \in L^2(\Omega)$ which can be approximated by smooth $C_c^\infty(\Omega')$ functions $\varphi_i \to \varphi$ in $L^2(\Omega)$. Since $\Omega' \subset\subset \Omega$ we also have $\varphi_i \to \varphi$ in $L^1(\Omega)$ (observe $\|\varphi\|_{L^1(\Omega)} = 1$ by construction of φ). Then

$$\fint_{B(x,r)} \partial_\ell u = \lim_{i\to\infty} \int_{B(x,r)} \partial_\ell u \varphi_i,$$

which leads to

$$|\fint_{B(x,r)} \partial_\ell u| \leq \Lambda \lim_{i\to\infty} \|\varphi_i\|_{L^1(\Omega')} \leq \Lambda \|\varphi\|_{L^1(\Omega')} = \Lambda.$$

Since x was chosen to be a Lebesgue point of $\partial_\ell u$, and since the last estimate holds for any $r > 0$, we find

$$|\partial_\ell u(x)| = \lim_{r\to 0} |\fint_{B(x,r)} \partial_\ell u| \leq \Lambda.$$

This again holds for any Lebesgue point $x \in \Omega$, and since almost all points in Ω are Lebesgue points,

$$|\partial_\ell u(x)| \leq \Lambda \quad \text{a.e. } x \in \Omega,$$

which implies

$$\|\partial_\ell u\|_{L^\infty(\Omega)} = \operatorname*{ess\,sup}_{\Omega} |\partial_\ell u| \leq \Lambda,$$

which was the claim. □

Theorem 13.24 ($C^{k-1,1} \approx W^{k,\infty}$). *(1) Let $\Omega \subset \mathbb{R}^n$ be open and nonempty, $k \in \mathbb{N}$ then*

$$C^{k-1,1}(\overline{\Omega}) \subset W^{k,\infty}(\Omega),$$

and the distributional derivative $D^k u$ belongs to L^∞ and we have

$$\|D^k u\|_{L^\infty(\Omega)} \leq C\,[D^{k-1}u]_{\mathrm{Lip}\,(\Omega)}$$

(2) Let $\Omega \subset\subset \mathbb{R}^n$ connected, $\partial\Omega \in C^{0,1}$. Then for $k \in \mathbb{N}$

$$W^{k,\infty}(\Omega) \subset C^{k-1,1}(\overline{\Omega}),$$

and

$$[D^{k-1}u]_{\mathrm{Lip}\,(\Omega)} \leq C(k,\Omega)\ \|D^k u\|_{L^\infty(\Omega)}.$$

The above holds in the following sense: recall that functions in L^p (and thus in particular in $W^{k,\infty}$ are classes (namely: two functions $f, g \in L^p(\Omega)$ are the same if they coincide almost everywhere). So what we mean above is: For every $f \in W^{k,\infty}(\Omega)$ there exists a representative $g \in C^{k-1,1}(\Omega)$ that coinides with f a.e.

Proof of Theorem 13.24. We restrict our attention to $k = 1$ and leave the other cases as exercise.

For (1): Let $u \in C^{0,1}(\overline{\Omega})$. Since u is Lipschitz,

$$\sup_{\Omega' \subset\subset \Omega}\ \sup_{|h| < \mathrm{dist}\,(\Omega', \partial\Omega)} \|\Delta_h^{e_\ell} u\|_{L^\infty(\Omega')} \leq [u]_{\mathrm{Lip}\,(\Omega)}.$$

From Proposition 13.23(2) we then obtain $u \in W^{1,\infty}(\Omega)$ with the claimed estimate.

For (2):

<u>First we assume that $\Omega = B(0,1)$ is a ball</u>, $u \in W^{1,\infty}(B(0,1))$. We argue by mollification (what else can we do): Let u_ε be the usual mollification $u_\varepsilon = \eta_\varepsilon * u$ which, as we already know, converges in $W^{1,2}_{loc}(B(0,1))$ to u. Moreover (also as seen before), for any $\delta \in (0,1)$, $x \in B(0,\delta)$, if $\varepsilon < \delta$ then

$$\partial_\ell u_\varepsilon(x) = \int \partial_\ell u(y)\ \eta_\varepsilon(y-x)\,dy,$$

and thus whenever $x \in B(0,\delta)$, if $\varepsilon < \delta$

$$|\partial_\ell u_\varepsilon(x)| \leq \|\partial_\ell u(y)\|_{L^\infty(B(0,1))} \|\eta_\varepsilon\|_{L^1(B(0,1))} = \|\partial_\ell u(y)\|_{L^\infty(B(0,1))}.$$

In particular, by the fundamental theorem of calculus (recall: u_ε is differentiable), whenever $\varepsilon < \delta$

$$[u_\varepsilon]_{\mathrm{Lip}\,,B(0,1-\delta)} \leq \|Du\|_{L^\infty(B(0,1))}.$$

Observe there is no constant on the right-hand side. Since moreover $\|u_\varepsilon\|_{L^\infty(B(0,\delta 0)} \leq \|u\|_{L^\infty(B(0,1))}$ we have that u_ε is equicontinuous and bounded, and thus by Arzela-Ascoli (up to a subsequence $\varepsilon \to 0$) we have $u_\varepsilon \to u$ in $C^0(B(0,\delta))$ (Here is where we find the "continuous representative of u", the limit of u_ε coinicides a.e. with u). In particular, u is continuous in $B(0,1-\delta)$. Also observe that for any $x \neq y \in B(0,1-\delta)$, for any $\varepsilon < \delta$,

$$|u(x)-u(y)| \leq 2\|u-u_\varepsilon\|_{L^\infty(B(0,1-\delta))} + |u_\varepsilon(x)-u_\varepsilon(y)| \leq 2\|u-u_\varepsilon\|_{L^\infty(B(0,1-\delta))} + |x-y|\|Du\|_{L^\infty(B(0,1))}.$$

This holds for any $\varepsilon < \delta$, so letting $\varepsilon \to 0$ we obtain by the uniform convergence $u_\varepsilon \to u$ in $B(0,1-\delta)$.

$$|u(x)-u(y)| \leq |x-y|\|Du\|_{L^\infty(B(0,1))} \quad \text{for all } x,y \in B(0,1-\delta)$$

This again holds for any $\delta > 0$ so that

$$|u(x)-u(y)| \leq |x-y|\|Du\|_{L^\infty(B(0,1))} \quad \text{for all } x,y \in B(0,1).$$

That is, u is Lipschitz continuous and we have

$$[u]_{\mathrm{Lip}\,(B(0,1))} \leq \|Du\|_{L^\infty(B(0,1))}.$$

So Theorem 13.24(2) is established for $\Omega = B(0,1)$.

Next assume that <u>$\Omega \subset\subset \mathbb{R}^n$ and $\partial\Omega \in C^{0,1}$. Moreover we assume Ω is path-connected.</u>

The regularity of the boundary is used in the following way: For any two points $x,y \in \Omega$ there exists a continuous path γ connecting x and y inside Ω such that the length of γ, $\mathcal{L}(\gamma) \leq C(\Omega)|x-y|$ (essentially take the straight line connecting x and y, when it hits $\partial\Omega$ follow $\partial\Omega$, then regularize and shift it away from $\partial\Omega$).

Since Ω is open, and by the argument above in every open ball we can replace u by its continuous representative we may assume that u w.l.o.g. is continous, and we just want to show that u is Lipschitz continuous.

Let $x, y \in \Omega$ and let γ be such a path connecting x and y. Set $\delta := \frac{1}{2}\mathrm{dist}\,(\gamma, \partial\Omega) > 0$. Setting $L := \lceil \frac{\mathcal{L}(\gamma)}{\delta} \rceil + 2$ points $(x_i)_{i=1}^L$ in γ, such that

$$\bigcup_{i=1}^{L} B(x_i, \delta) \supset \gamma.$$

and such that $B(x_i, \delta) \cap B(x_{i+1}, \delta) \neq \emptyset$, $x \in B(x_1, \delta)$ and $y \in B(x_L, \delta)$. In every $B(x_i, \delta)$ we use the argument from above, and have

$$[u]_{\mathrm{Lip}\,(B(x_i,\delta))} \le \|Du\|_{L^\infty(B(x_i,\delta))} \le \|Du\|_{L^\infty(\Omega)}.$$

Now, by triangular inequality

$$\begin{aligned}|u(x) - u(y)| \le & |u(x) - u(x_0)| + |u(y) - u(x_L)| + \sum_{i=1}^{L-1} |u(x_i) - u(x_{i+1})| \\ \le & \|Du\|_{L^\infty(\Omega)}\,(L+1)2\delta \le C\mathcal{L}(\gamma) \le C(\Omega)\|Du\|_{L^\infty(\Omega)}\,|x-y|.\end{aligned}$$

This implies that u is Lipschitz continuous with

$$[u]_{\mathrm{Lip}} \le C(\Omega)\|Du\|_{L^\infty(\Omega)}.$$

which establishes the theorem. $\square$

13.3. **Remark: Weak compactness in $W^{k,p}$.** In the proof of of Proposition 13.23(2) we derived and used the following consequence of Theorem 12.13, see also Proposition 12.16, which we want to record (so we don't have to argue always with Theorem 12.13.

Theorem 13.25 (Weak compactness)**.** *Let $1 < p < \infty$, $k \in \mathbb{N}$, $\Omega \subset \mathbb{R}^n$ open. Assume that $(f_i)_{i\in\mathbb{N}}$ is a bounded sequence in $W^{k,p}(\Omega)$, that is*

$$\sup_{i\in\mathbb{N}} \|f_i\|_{W^{k,p}(\Omega)} < \infty.$$

Then there exists a function $f \in W^{k,p}(\Omega)$ and a subsequence f_{i_j} such that f_{i_j} weakly $W^{k,p}$-converges to f, that is for any $|\gamma| \le k$ and any $g \in L^{p'}(\Omega)$, where $p' = \frac{p}{p-1}$ is the Hölder dual of p, we have

$$\int_\Omega \partial^\gamma f_{i_j}\, g \xrightarrow{i\to\infty} \int_\Omega \partial^\gamma f\, g.$$

In particular we have

$$\|f\|_{W^{k,p}(\Omega)} \le \limsup_i \|f_i\|_{W^{k,p}(\Omega)}.$$

To obtain this statement one either proves it by hand from the L^p-version. Or, one considers $W^{1,p}$ as a closed subspace of $L^p \times \ldots L^p$, as in the proof of Corollary 10.12 and used that *convex closed subsets are weakly closed*, Theorem 12.26.

13.4. **Extension Theorems.** If f is a Lipschitz function on a set $\Omega \subset \mathbb{R}^n$, then f can be thought of as a restriction of a map $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$, $f = \tilde{f}\Big|_{\Omega}$. This is (a special case of) the so-called Kirszbraun theorem. This is in general not true for Sobolev functions, even if Ω is open.

Definition 13.26. Let $\Omega \subset \mathbb{R}^n$ be open. Ω is called a *$W^{k,p}$-extension domain*, if there exists a linear operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^n)$ such that

$$Eu(x) = u(x) \quad \text{for all } x \in \mathbb{R}^n,\ u \in W^{k,p}(\Omega)$$

and E is bounded, i.e.

$$\sup_{\|u\|_{W^{k,p}(\Omega)} \le 1} \|Eu\|_{W^{k,p}(\mathbb{R}^n)} < \infty.$$

Theorem 13.27. *Any open set $\Omega \subset\subset \mathbb{R}^n$ with boundary $\partial\Omega \in C^k$ is a $W^{k,p}(\Omega)$ extension domain for $k \in \mathbb{N}$, $1 \le p < \infty$.*

More precisely, for any $\tilde{\Omega} \supset\supset \Omega$ there exists an operator $E : W^{k,p}(\Omega) \to W^{k,p}_0(\tilde{\Omega})$ with $Eu = u$ in Ω and

$$\|Eu\|_{W^{k,p}(\tilde{\Omega})} \le C(\Omega, \tilde{\Omega}, n, k)\, \|u\|_{W^{k,p}(\Omega)}.$$

Remark 13.28. Theorem 13.27 is not optimal w.r.t to the regularity of $\partial\Omega$. Indeed one can show that any Lipschitz domain, i.e. any $\Omega \subset \mathbb{R}^n$ open with $\partial\Omega$ locally a Lipschitz-graph, is a $W^{k,p}$-extension domain. For non-Lipschitz-domains this may not be true, take e.g. the example in Exercise 13.30.

Proof. We will first show how to extend $W^{k,p}$-functions from $\mathbb{R}^n_+$ to all of $\mathbb{R}^n$. Then by "flattening the boundary" (for this we need the regularity $\partial\Omega$) we extend this argument to general Ω as claimed.

<u>From $\mathbb{R}^n_+$ to $\mathbb{R}^n$:</u>

Denote the variables in $\mathbb{R}^n$ by (x', x_n) where $x' \in \mathbb{R}^{n-1}$ and $x_n \in \mathbb{R}$.

We can explicitely define $E_0 : W^{k,p}(\mathbb{R}^n_+) \to W^{k,p}(\mathbb{R}^n)$ by a type of reflection.

The main point is that we know (from the heaviside function example) that $W^{1,k}$-functions cannot have a jump, so at least for smooth functions u, if we hope for $E_0 u \in W^{1,p}$ we need that

$$\lim_{y_n \to 0^-} E_0 u(y', y_n) \overset{!}{=} \lim_{y_n \to 0^+} E_0 u(y', y_n) = \lim_{y_n \to 0^+} u(y', y_n)$$

So, for $k = 1$ we could simply use the even reflection,

$$E_0 u(y', y_n) := u(y', |y_n|) = \begin{cases} u(y', y_n) & \text{if } y_n > 0 \\ u(y, -y_n) & \text{if } y_n < 0. \end{cases}$$

which indeed takes $C^\infty(\overline{\mathbb{R}^n_+})$-functions into Lipschitz-functions (i.e. $W^{1,\infty}_{loc}(\mathbb{R}^n)$-functions, hence $W^{1,p}_{loc}(\mathbb{R}^n)$).

More generally, for $W^{k,p}$-functions, $k \geq 1$ we then need that for any $\ell = 1,\ldots,k$ the $(\ell-1)$-th derivatives in y_n-direction coincide:

$$\lim_{y_n\to 0^-} (\partial_n)^{\ell-1} E_0 u(y',y_n) \overset{!}{=} \lim_{y_n\to 0^+} (\partial_n)^{\ell-1} E_0 u(y',y_n) = \lim_{y_n\to 0^+} (\partial_n)^{\ell-1} u(y',y_n). \tag{13.11}$$

So again, we use a reflection, but a more complicated one,

$$E_0 u(y',y_n) := \begin{cases} u(y',y_n) & \text{if } y_n > 0\\ \sum_{i=1}^k \sigma_i\, u(y,-iy_n) & \text{if } y_n < 0.\end{cases}$$

Here, $(\sigma_i)_{i=1}^k$ are constants to be chosen, such that (13.11) is true for smooth functions: For all $\ell = 1,\ldots,k$

$$\sum_{i=1}^k \sigma_i (-i)^{\ell-1} (\partial_n)^\ell u(x',0) = (\partial_n)^\ell u(x',0) \quad \Leftarrow \quad \sum_{i=1}^k \sigma_i (-i)^{\ell-1} = 1.$$

Such a σ exists by linear algebra: Defining a matrix A by $A_{i\ell} := (-i)^{\ell-1}$, and interpreting σ as a vector in $\mathbb{R}^k$ we want to solve

$$A\sigma = \begin{pmatrix} 1\\ \vdots\\ 1\end{pmatrix},$$

which is possible if A is invertible (to check this is the case is left as an exercise).

Now we argue as follows: Let $u \in W^{k,p}(\mathbb{R}^n_+)$. By Proposition 13.18 there exists $u_j \in C_c^\infty(\mathbb{R}^{n-1}\times[0,\infty))$ that approximate u in $W^{k,p}(\mathbb{R}^n_+)$.

One now checks that $E_0 u_j \in C^{k-1,1}(\mathbb{R}^n)$, moreover we have for almost any $x \in \mathbb{R}^n$ (namely whenever $x = (x',x_n)$, with $x_n \neq 0$), for any $|\gamma| \leq k$,

$$|\partial^\gamma (E_0 u_j)(x',x_n)| \leq C(\sigma,k) \begin{cases} |\partial^\gamma u_j(x',x_n)| & x_n > 0\\ \sum_{i=1}^k |\partial^\gamma u(y,-iy_n)| & x_n < 0.\end{cases}$$

Let us illustrate this fact for $k = 1$, for $k > 1$ it is an exercise.

$$\partial_{x_\alpha}(E_0 u_j)(x',x_n) = \partial_{x_\alpha} u_j(x',|x_n|) = (\partial_{x_\alpha} u_j)(x',|x_n|) \quad \alpha = 1,\ldots,n-1.$$

$$\partial_{x_n}(E_0 u_j)(x',x_n) = \partial_{x_\alpha} u_j(x',|x_n|) = (\partial_{x_n} u_j)(x',|x_n|)\frac{x_n}{|x_n|}.$$

In particular we get that $u \in W^{k,p}(\mathbb{R}^n)$ and

$$\|E_0 u_j\|_{W^{k,p}(\mathbb{R}^n)} \leq C(k)\ \|u_j\|_{W^{k,p}(\mathbb{R}^n_+)}$$

In particular we get

$$\limsup_{j\to\infty} \|E_0 u_j\|_{W^{k,p}(\mathbb{R}^n)} \leq C(k)\ \limsup_{j\to\infty} \|u_j\|_{W^{k,p}(\mathbb{R}^n_+)} = C(k)\ \|u\|_{W^{k,p}(\mathbb{R}^n_+)}.$$

Thus, in view of Theorem 13.25 we find $g \in W^{k,p}$ which is the weak $W^{k,p}$-limit of $E_0 u_j$. By strong L^p-convergence of u_j to u we see that indeed $E_0 g = u$, and thus we get

$$\|E_0 u\|_{W^{k,p}(\mathbb{R}^n)} \leq C(k)\ \|u\|_{W^{k,p}(\mathbb{R}^n_+)}.$$

as claimed.

From Ω to $\mathbb{R}^n$ We only sketch the remaining arguments. If $\partial\Omega \in C^k$ then from small balls B centered at boundary points there exists C^k-charts $\phi : B \to \mathbb{R}^n$ such that $\phi(B \cap \Omega) \subset \mathbb{R}^n_+$ $\phi(B \cap \Omega^c) \subset \mathbb{R}^n_-$. By a decomposition of unity, we set $u = \sum_i \eta_i u$ such that η_i are supported only in one of these balls $B_i \cap \Omega$. Then $(\eta_i u) \circ \phi_i \in W^{k,p}(\mathbb{R}^n_+)$ (since it is locally in $W^{k,p}$ and then it is constantly zero). Here we use that (we haven't shown it) the Transformation rule still holds for Sobolev functions. Then we extend $(\eta_i \circ u) \circ \phi_i$ to all of $\mathbb{R}^n$, i.e. consider$E_0((\eta_i u) \circ \phi_i)$. Finally we set

$$E_1 u := \sum_i \left(E_0((\eta_i u) \circ \phi_i)\right) \circ \phi_i^{-1}.$$

The transformation rule shows that $Eu \in W^{k,p}(\mathbb{R}^n)$.

From Ω to Ω' To get $E_2 u \in W^{k,p}_0(\Omega')$ we simply take a cuttoff function $\eta \in C^\infty_c(\Omega')$, $\eta \equiv 1$ in Ω, and set

$$E_2 u := \eta E_1 u.$$

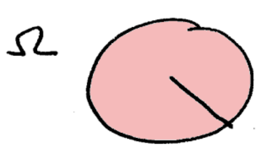
□

Whenever Ω is a $W^{k,p}$-extension domain, smooth $C^\infty(\overline{\Omega})$-functions are dense in $W^{k,p}(\Omega)$. In particular in view of Remark 13.28 the approximation holds for Lipschitz domains.

Exercise 13.29. *Assume Ω is a $W^{k,p}$-extension domain, $1 \leq p < \infty$. Show that any $u \in W^{k,p}(\Omega)$ can be approximated by smooth functions in $C^\infty(\overline{\Omega})$ (essentially: prove Theorem 13.15 for such Ω).*

Observe that for general sets, $C^\infty(\Omega)$ (not $C^\infty(\overline{\Omega})$) is dense in $W^{k,p}(\Omega)$, Theorem 13.16.

Exercise 13.30. *Let Ω be a two-dimensional disc with one radius removed (see picture below), e.g. $\Omega = B(0,1) \backslash [0,1] \times \{0\}$. Prove that $C^\infty(\overline{\Omega})$ functions are not dense in $W^{1,p}(\Omega)$.*



Why does this not contradict Theorem 13.15 or Theorem 13.27?

13.5. **Traces.** Let $\Omega \subset \mathbb{R}^n$ be *open* and (for simplicity) $\partial\Omega \in C^\infty$.

If $u \in C^\alpha(\Omega)$, $\alpha \in (0,1]$ with

$$\|u\|_{C^\alpha(\Omega)} := \sup_{x\in\Omega} |u(x)| + \sup_{x\neq y\in\Omega} \frac{|u(x)-u(y)|}{|x-y|^\alpha} < \infty$$

then we find a unique map $u\Big|_{\partial\Omega} \in C^\alpha(\partial\Omega)$. Indeed, for any $\overline{x} \in \partial\Omega$ there exists exactly one value $u(\overline{x})$ such that $\overline{u}(\overline{x}) = \lim_{\Omega\ni x\to\overline{x}} u(x)$, because $\|u\|_{L^\infty} < \infty$ and since we have uniform continuity

$$|u(x)-u(y)| \leq \|u\|_{C^\alpha(\Omega)}|x-y|^\alpha \xrightarrow{|x-y|\to 0} 0.$$

Moreover, for any $\overline{x}, \overline{y} \in \partial\Omega$ and $x, y \in \Omega$ we have

$$\begin{aligned}|\overline{u}(\overline{x}) - \overline{u}(\overline{y})| \leq& |\overline{u}(\overline{x}) - u(x)| + |u(x)-u(y)| + |u(x) - \overline{u}(\overline{y})|\\ \leq& |\overline{u}(\overline{x}) - u(x)| + \|u\|_{C^\alpha}|x-y|^\alpha + |u(x) - \overline{u}(\overline{y})|\end{aligned}$$

Taking $x \to \overline{x}$ and $y \to \overline{y}$ we thus find

$$|\overline{u}(\overline{x}) - \overline{u}(\overline{y})| \leq \|u\|_{C^\alpha}|\overline{x}-\overline{y}|^\alpha$$

that is

$$\|\overline{u}\|_{C^\alpha(\partial\Omega)} \leq \|u\|_{C^\alpha(\Omega)}.$$

The map that computes from u the trace map $\overline{u}$ we may call T, $\overline{u} = Tu$. Then we have a linear operator

$$T : C^{k,\alpha}(\overline{\Omega}) \to C^{k,\alpha}(\partial\Omega),$$

By the computations above, T is linear and bounded

$$\|Tu\|_{C^{k,\alpha}(\partial\Omega)} \leq \|u\|_{C^{k,\alpha}(\Omega)}.$$

On the other hand, when $u \in L^p(\Omega)$ there is absolutely no reasonable (unique) sense of a trace $u\Big|_{\partial\Omega}$.

One interesting and important fact of Sobolev spaces is that there is such a trace operator T if $k - \frac{1}{p} > 0$, that associates to a Sobolev function $u \in W^{k,p}(\Omega)$ a map $Tu \in W^{k-\frac{1}{p},p}(\partial\Omega)$. Observe that formally, if $p = \infty$ (i.e. in the Lipschitz case, the trace map is of the same class as the interior map, but for $p < \infty$ the trace map has less differentiability than the interior map. We do not want to deal with fractional Sobolev spaces here, so instead of proving the sharp trace estimate

$$T : W^{1,p}(\Omega) \to W^{1-\frac{1}{p},p}(\partial\Omega)$$

we will only show the following:

Theorem 13.31. *Let $\Omega \subset\subset \mathbb{R}^n$, $\partial\Omega \in C^1$, $1 \leq p < \infty$. There exists a (unique) bounded and linear Trace operator T*

$$T : W^{1,p}(\Omega) \to L^p(\partial\Omega)$$

such that

(1) $Tu = u\Big|_{\partial\Omega}$ *whenever* $u \in C^0(\overline{\Omega}) \cap W^{1,p}(\Omega)$.
(2) for each $u \in W^{1,p}(\Omega)$ *we have*

$$\|Tu\|_{L^p(\partial\Omega, d\mathcal{H}^{n-1})} \leq C(\Omega,p)\ \|u\|_{W^{1,p}(\Omega)}.$$

Proof. For $u \in C^1(\overline{\Omega}) \cap W^{1,p}(\Omega)$ we define

$$Tu := u\Big|_{\partial\Omega}.$$

It now suffices to show that for all $u \in C^1(\overline{\Omega}) \cap W^{1,p}(\Omega)$ we have

$$\|Tu\|_{L^p(\partial\Omega)} \leq C(\Omega,p)\ \|u\|_{W^{1,p}(\Omega)}. \tag{13.12}$$

Then, by density of smooth functions $C^\infty(\overline{\Omega})$ in $W^{1,p}(\Omega)$, Theorem 13.15, linearity and boundedness of the trace operator, there exists a (unique) extension of T to all of $W^{1,p}(\Omega)$.

To see (13.12) we argue again first on a flat boundary $\Omega = \mathbb{R}^n_+$. A flattening the boundary argument as above, then leads to the claim.

Observe the following, which holds by the integration-by-parts formula:

$$\|u\|^p_{L^p(\mathbb{R}^{n-1})} = \int_{\mathbb{R}^{n-1}\times\{0\}} |u(x')|^p d\mathcal{H}^{n-1}(x) = \int_{\mathbb{R}^n_+} \partial_n\left(|u(x)|^p\right) dx = \int_{\mathbb{R}^n_+} p|u(x)|^{p-2}u(x)\,\partial_n u(x)\,dx$$

Then by Young's inequality, $ab \leq C(a^p + b^{p'})$ (where $p' = \frac{p}{p-1}$ is the Hölder dual of p),

$$\int_{\mathbb{R}^n_+} p|u(x)|^{p-2}u(x)\,\partial_n u(x)\,dx \leq C\ \int_{\mathbb{R}^n_+} \left(|u|^p + |\partial_n u|^p\right) \leq C\|u\|^p_{W^{1,p}(\mathbb{R}^n_+)}.$$

(this still works if $p=1$ and $p'=\infty$!). This establishes (13.12) for $\Omega = \mathbb{R}^n_+$. For general Ω we use a decomposition of unity and flattening the boundary argument as in the theorems above. $\square$

Theorem 13.32 (Zero-boundary data and traces)**.** *Let* $\Omega \subset\subset \mathbb{R}^n$ *and* $\partial\Omega \in C^1$. *Let* $u \in W^{1,p}(\Omega)$.

Then $u \in H^{1,p}_0(\Omega)$ *is equivalent to* $u \in W^{1,p}_0(\Omega)$, *where*

$$W^{1,p}_0(\Omega) = \{u \in W^{1,p}(\Omega) : Tu = 0\}$$

for the trace operator T *from Theorem 13.31.*

Remark 13.33. By induction one obtains that if $\partial\Omega \in C^\infty$ then $H^{k,p}_0(\Omega)$ are exactly those functions where $T(\partial^\gamma u) = 0$ for any $|\gamma| \leq k-1$.

For time reasons we will not give the proof here. For a proof see [Evans, 2010, §5.5, Theorem 2].

Remark 13.34. The trace theory can be extended to Lipschitz maps and improved on the boundary.

More precisely, let $\Omega \subset \mathbb{R}^n$ be an open set (for simplicity: bounded) such that $\partial\Omega$ is locally a Lipschitz Graph (for simplicity: compact).

Denote for $s \in (0,1)$ the *fractional Sobolev space* (also called Gagliardo/Slobodeckij space, there are many other fractional Sobolev space)

$$W^{s,p}(\partial\Omega) := \left\{ f \in L^p(\partial\Omega) : \quad [f]_{W^{s,p}(\partial\Omega)} < \infty \right\},$$

where (the integral on the boundary is the $\mathcal{H}^{n-1}$-integral)

$$[f]_{W^{s,p}(\partial\Omega)} := \left(\int_{\partial\Omega} \int_{\partial\Omega} \frac{|f(x) - f(y)|^p}{|x-y|^{n-1+sp}} \, dx \, dy \right)^{\frac{1}{p}}$$

One can show that $W^{s,p}(\partial\Omega)$ is a Banach space when equipped with the norm

$$\|f\|_{W^{s,p}(\partial\Omega)} := \|f\|_{L^p(\partial\Omega)} + [f]_{W^{s,p}(\partial\Omega)}.$$

Then there exists a trace map $T : W^{1,p}(\Omega) \to W^{1-\frac{1}{p},p}(\partial\Omega)$ (if $p = \infty$: $W^{1,\infty}$ to Lip) with the following properties

- T is linear and continuous, i.e.
$$\|Tf\|_{W^{1-\frac{1}{p},p}(\partial\Omega)} \le C(\Omega) \, \|f\|_{W^{1,p}(\Omega)}$$
- If $f \in W^{1,p}(\Omega) \cap C^0(\overline{\Omega})$ then $Tf = f\Big|_{\partial\Omega}$
- There exists a linear bounded operator $S : W^{1-\frac{1}{p},p}(\partial\Omega) \to W^{1,p}(\Omega)$,
$$\|Sf\|_{W^{1,p}(\Omega)} \le \tilde{C}(\Omega) \, \|f\|_{W^{1-\frac{1}{p},p}(\partial\Omega)}$$
  such that $T \circ S = \mathrm{id}$.

This theorem is due to Gagliardo, [Gagliardo, 1957], see also [Mironescu, 2005].

13.6. **Embedding theorems.** Let X, Y be two Banach spaces. $T : X \to Y$ is a (we assume always: linear) embedding if T is injective. We say that the *embedding* $X \subset Y$ is continuous under the operator T, if T is a linear embedding and T is *continuous* (i.e. a bounded operator). If (as it often happens) T is (in a reasonable sense) the identity map, then we say that X embedds into Y continuously, and write $X \hookrightarrow Y$. E.g., clearly (by definition)

$$W^{1,p}(\Omega) \hookrightarrow L^p(\Omega)$$

since, by definition of the norm

$$\|u\|_{L^p(\Omega)} \le \|u\|_{W^{1,p}(\Omega)}.$$

Recall that we say than an embedding $X \hookrightarrow Y$ is *compact* if the operator $T : X \to Y$ is compact, i.e. if for any bounded sequence $(x_n)_{n\in\mathbb{N}} \subset X$, $\sup_n \|x_n\|_X < \infty$, we have that $(T(x_n))_{n\in\mathbb{N}} \subset Y$ has a convergent subsequence in Y.

From functional analysis we also have: If $T : X \to Y$ is compact, then if $(x_n)_{n\in\mathbb{N}}$ is weakly convergent in X then Tx_n is strongly convergent in Y.

By Arzela-Ascoli, it is easy to check that $C^{k,\alpha}(\overline{\Omega})$ embedds compactly into $C^{\ell,\beta}(\overline{\Omega})$ if $k \geq \ell$ and $k + \alpha > \ell + \beta$.

The first important theorem is that for bounded sets Ω with smooth boundary we have $W^{1,p}(\Omega)$ embedds compactly into $L^p(\Omega)$. (By induction: $W^{k,p}(\Omega)$ embedds compactly into $W^{\ell,p}(\Omega)$ whenever $k \geq p$).

Observe that by Theorem 13.25 we have weak compactness in $W^{1,p}(\Omega)$ for bounded set, but strong convergence in $L^p(\Omega)$.

Theorem 13.35 (Rellich-Kondrachov). *Let $\Omega \subset\subset \mathbb{R}^n$, $\partial\Omega \subset C^{0,1}$, $1 \leq p \leq \infty$. Assume that $(u_k)_{k\in\mathbb{N}} \in W^{1,p}(\Omega)$ is bounded, i.e.*

$$\sup_{k\in\mathbb{N}} \|u_k\|_{W^{1,p}(\Omega)} < \infty.$$

Then there exists a subsequence $k_i \to \infty$ and $u \in L^p(\Omega)$ such that u_{k_i} is (strongly) convergent in $L^p(\Omega)$, moreover the convergence is pointwise a.e..

Proof. If $p = \infty$, from Theorem 13.24 we have $W^{1,\infty}(\Omega) = C^{0,1}(\overline{\Omega})$. By Arzela-Ascoli it is clear that $C^{0,1}$ is compactly embedded in $C^0(\overline{\Omega})$, so in particular in L^∞ (which has the same norm as $C^0(\overline{\Omega})$.

Now let $p \in [1, \infty)$. By the extension theorem, Theorem 13.27 we may assume that $u_k \in W^{1,p}(\mathbb{R}^n)$ with $\operatorname{supp} u_k \subset B(0, R)$ for some (fixed) large $R > 0$.

The main idea is to use Arzela-Ascoli for mollified versions of u_k. Denote by $\eta \in C_c^\infty(B(0, 1))$ the usual bump function, $\int \eta = 1$, and $\eta_\varepsilon = \varepsilon^{-n}\eta(\cdot/\varepsilon)$. Set

$$u_{k,\varepsilon} := \eta_\varepsilon * u_k \in C_c^\infty(B(0, 2R)).$$

Observe that

$$|u_{k,\varepsilon}(x)| \leq C(R)\varepsilon^{-n}\|u_k\|_{L^p(B(0,R))}$$

$$|Du_{k,\varepsilon}(x)| \leq C(R)\varepsilon^{-n-1}\|u_k\|_{L^p(B(0,R))},$$

so since u_k is bounded (even L^p-boundedness is enough for now) we have

$$\sup_{k\in\mathbb{N}} \|u_{k,\varepsilon}\|_{\operatorname{Lip}(\mathbb{R}^n)} \leq C(\varepsilon).$$

That is, for any $\varepsilon_j := \frac{1}{j}$ there exists a subsequence $u_{k_{i(\varepsilon_j)},\varepsilon_j}$ that is convergent in $L^\infty(\mathbb{R}^n)$.

By a diagonalizing this subsequences we obtain only one subsequence u_{k_i,ε_j} so that for any fixed ε_j we have convergence in $L^\infty(\mathbb{R}^n)$, i.e. for any $j \in \mathbb{N}$ and any $\delta > 0$ there exists $N_{j,\delta} \in \mathbb{N}$ such that

$$\|u_{k_{i_1},\varepsilon_j} - u_{k_{i_2},\varepsilon_j}\|_{L^\infty(\mathbb{R}^n)} \leq \delta \quad \forall i_1, i_2 > N_{j,\delta}.$$

Next we observe, by the fundamental theorem of Calculus,

$$\begin{aligned}\left|u_{k_{i_1}}(x) - u_{k_{i_1},\varepsilon_j}(x)\right| =& \int_{\mathbb{R}^n} |\eta_\varepsilon(z)| \left|u_{k_{i_1}}(x-z) - u_{k_{i_1}}(x)\right| \leq \int_0^1 \int_{B(0,\varepsilon)} |\eta_\varepsilon(z)||Du_{k_{i_1}}(x-tz))|\,|z|dz\,dt \\ \leq& \varepsilon^{1-n} \int_0^1 \left(\int_{B(0,\varepsilon)} |Du_{k_{i_1}}(x-tz))|^p\right)^{\frac{1}{p}} \varepsilon^{n-\frac{n}{p}} dz\,dt\end{aligned}$$

Thus, by Fubini

(13.13)

$$\begin{aligned}\|u_{k_{i_1}} - u_{k_{i_1},\varepsilon_j}\|_{L^p(\mathbb{R}^n)} =& \|u_{k_{i_1}} - u_{k_{i_1},\varepsilon_j}\|_{L^p(B(0,2R))} \leq \varepsilon^{1-\frac{n}{p}} \left(\int_0^1 \int_{B(0,R)} \int_{B(0,\varepsilon)} |Du_{k_{i_1}}(x-tz) dxdzdt\right)^{\frac{1}{p}} \\ \leq& \varepsilon \|Du_{k_{i_1}}\|_{L^p(B(0,2R))}.\end{aligned}$$

Now we claim that this leads to a Cauchy-sequence for the (non-mollified) u_{k_i}: Let $\delta > 0$.

$$\begin{aligned}\|u_{k_{i_1}} - u_{k_{i_2}}\|_{L^p(\mathbb{R}^n)} \leq& \|u_{k_{i_1}} - u_{k_{i_1},\varepsilon_j}\|_{L^p(\mathbb{R}^n)} + \|u_{k_{i_1},\varepsilon_j} - u_{k_{i_2},\varepsilon_j}\|_{L^p(\mathbb{R}^n)} + \|u_{k_{i_2},\varepsilon_j} - u_{k_{i_2}}\|_{L^p(\mathbb{R}^n)} \\ \overset{(13.13)}{\leq}& 2\,C\,\varepsilon_j \sup_k \|u_k\|_{W^{1,p}(\mathbb{R}^n)} + C(R)\|u_{k_{i_1},\varepsilon_j} - u_{k_{i_2},\varepsilon_j}\|_{L^\infty(B(2R))}\end{aligned}$$

Choosing now first ε_j small enough so that

$$2\,C\,\varepsilon_j \sup_k \|u_k\|_{W^{1,p}(\mathbb{R}^n)} < \frac{\delta}{2}$$

and then choosing for this ε_j the $N(\varepsilon_j, \delta)$ large enough so that for any $i_1, i_2 > N(\varepsilon_j, \delta)$

$$C(R)\|u_{k_{i_1},\varepsilon_j} - u_{k_{i_2},\varepsilon_j}\|_{L^\infty(B(2R))} < \frac{\delta}{2}$$

we see that

$$\|u_{k_{i_1}} - u_{k_{i_2}}\|_{L^p(\mathbb{R}^n)} \leq \delta \quad \text{for any } i_1, i_2 > N(\varepsilon_j, \delta).$$

That is, u_{k_i} is a Cauchy sequence in $L^p(\mathbb{R}^n)$ and thus converges. □

One important consequence of Rellich's theorem, Theorem 13.35 is Poincarè's inequality. In 1D it is called sometimes Wirtinger's inequality – and it is quite easy to prove. Let $I = (a,b) \subset \mathbb{R}$, then for any $u \in W^{1,p}(I)$,

$$\|u - (u)_I\|_{L^p(I)} \leq C(I,p)\|u'\|_{L^p(I)}. \tag{13.14}$$

Here

$$(u)_I := \fint_I u$$

denotes the mean value of u on I.

The proof of (13.14) is done by the fundamental theorem of calculus, Lemma 13.22. We have (using Hölder's inequality and Fubini many times)

$$\|u-(u)_I\|_{L^p(I)}^p \leq |I|^{-1}\int_I\int_I |u(x)-u(y)|^p \leq |I|^{-1}\int_0^1\int_I\int_I |u'(tx+(1-t)y)|^p\,|x-y|^p\,dx\,dydt$$

Now observe that by substituting $\tilde{y} := tx+(1-t)y$, we have

$$\begin{aligned}
&|I|^{-1}\int_0^{\frac{1}{2}}\int_I\int_I |u'(tx+(1-t)y)|^p\,|x-y|^p\,dx\,dydt\\
\leq & C(I,p)\int_0^{\frac{1}{2}}\int_I\int_I \frac{1}{1-t}|u'(\tilde{y})|^p\,d\tilde{y}\,dxdt\\
= & C(I,p)\int_0^{\frac{1}{2}}\frac{1}{1-t}\,dt\ |I|\ \|u'\|_{L^p(I)}^p\\
= & \tilde{C}(I,p)\|u'\|_{L^p(I)}^p.
\end{aligned}$$

In the same way, substituting $\tilde{x} := tx+(1-t)y$

$$\begin{aligned}
&|I|^{-1}\int_{\frac{1}{2}}^1\int_I\int_I |u'(tx+(1-t)y)|^p\,|x-y|^p\,dx\,dydt\\
\leq & \tilde{C}(I,p)\|u'\|_{L^p(I)}^p.
\end{aligned}$$

So (13.14) is established.

The Poincarè inequality says that (13.14) holds also in higher dimensions,

$$\|u-(u)_\Omega\|_{L^p(\Omega)} \leq C(\Omega,p)\|\nabla u\|_{L^p(\Omega)}. \tag{13.15}$$

If Ω is convex, the above proof works almost verbatim, in general open sets Ω this is more tricky.

Clearly, (13.15) does not hold if we remove $(u)_\Omega$ from the left-hand side. Indeed, just take $u \equiv const$ to find a counterexample. And indeed, a $W^{1,p}$-Poincaré-type inequality holds whenever *constants are excluded* in a reasonable sense.

Theorem 13.36 (Poincaré)**.** *Let $\Omega \subset\subset \mathbb{R}^n$ be open and connected, $\partial\Omega \in C^{0,1}$, $1 \leq p \leq \infty$.*

Let $K \subset W^{1,p}(\Omega)$ be a closed (with respect to the $W^{1,p}$-norm) cone with only one constant function $u \equiv 0$. That is, let $K \subset W^{1,p}(\Omega)$ be a closed set such that

(1) $u \in K$ implies $\lambda u \in K$ for any $\lambda \geq 0$.
(2) if $u \in K$ and $u \equiv const$ then $u \equiv 0$.

Then there exists a constant $C = C(K,\Omega)$ such that

$$\|u\|_{L^p(\Omega)} \leq C\|\nabla u\|_{L^p(\Omega)} \quad \forall u \in K. \tag{13.16}$$

The proof is a standard method from analysis, called a *blow-up proof*. One assumes that the claim is false, and then tries to compute/construct the "most extreme" counterexample

– which one then hopes to see cannot exist. Before we begin, we need the following small Lemma.

Lemma 13.37. *For $\Omega \subset \mathbb{R}^n$ open assume that $u \in W^{1,1}_{loc}(\Omega)$. If $\nabla u \equiv 0$ then u is constant in every connected component of Ω.*

Proof. This follows from (local) approximation by smooth functions. If $\nabla u \equiv 0$ then $\nabla u_\varepsilon \equiv 0$ in $\Omega_{-\varepsilon}$, where $u_\varepsilon = \eta_\varepsilon * u$. This implies that $u_\varepsilon \equiv const$ in every connected component of $\Omega_{-\varepsilon}$. Pointwise a.e. convergence of u_ε to u gives the claim. □

Proof of Theorem 13.36. Assume the claim is false for a given K as above. That means however we choose the constant C there will be some countexample u that dails the claimed inequality (13.16).

That is, for any $m \in \mathbb{N}$ there exists $u_m \in K$ such that (13.16) is false for $C = m$, i.e.

$$\|u_m\|_{L^p(\Omega)} > m \, \|\nabla u_m\|_{L^p(\Omega)}. \tag{13.17}$$

Now we construct the "extreme/blown up" counterexample (that, as we shall see, does not exist – leading to a contradicition).

Firstly, we can assume w.l.o.g.

$$\|u_m\|_{L^p(\Omega)} = 1, \quad \|\nabla u_m\|_{L^p(\Omega)} \leq \frac{1}{m}. \tag{13.18}$$

Indeed otherwise we can just take $\tilde{u}_m := \frac{u_m}{\|u_m\|_{L^p(\Omega)}}$ which satisfies (13.18).

(13.18) implies in particular,

$$\sup_m \|u_m\|_{W^{1,p}(\Omega)} < \infty.$$

In view of Rellich's theorem, Theorem 13.35, we can thus assume w.l.o.g. (otherwise taking a subsequence) that u_m is convergent in $L^p(\Omega)$. In particular u_m is a Cauchy sequence in $L^p(\Omega)$. Observe that also ∇u_m is a cauchy sequence in $L^p(\Omega)$, indeed by (13.18) $\nabla u_m \xrightarrow{m\to\infty} 0$ in $L^p(\Omega)$. In particular, u is a Cauchy sequence in $W^{1,p}(\Omega)$. Since $W^{1,p}(\Omega)$ is a Banach space we find a limit map $u \in W^{1,p}(\Omega)$ such that

$$\|u_m - u\|_{W^{1,p}(\Omega)} \xrightarrow{m\to\infty} 0. \tag{13.19}$$

In view of (13.18) this implies that $\nabla u \equiv 0$. From Lemma 13.37 and since Ω is connected, u is a constant map. But since K is closed we have that $u \in K$, and since the only constant map in K is the constant zero map, we find $u \equiv 0$ in Ω. But then by (13.19)

$$\|u_m\|_{W^{1,p}(\Omega)} \xrightarrow{m\to\infty} 0.$$

which contradicts the conditions in (13.18), namely

$$\|u_m\|_{W^{1,p}(\Omega)} \geq \|u_m\|_{L^p(\Omega)} \overset{(13.18)}{=} 1.$$

We have found a contradiction, and thus the assumption above (that for any m there exists u_m that contradicts the claimed equation) is false. So there must be some number m such that for $C := m$ the equation (13.16) holds. $\square$

Observe we have no idea what kind of $C = C(K)$ we get in Theorem 13.36 – which is somewhat the unsatisfying part of this type of blowup proof.

Corollary 13.38 (Poincaré type lemma)**.** *Let $\Omega \subset\subset \mathbb{R}^n$ be open, connected, and $\partial\Omega \in C^{0,1}$.*

(1) There exists $C = C(\Omega)$ such that for all $u \in W^{1,p}(\Omega)$ we have
$$\|u - (u)_\Omega\|_{L^p(\Omega)} \leq C(\Omega)\|\nabla u\|_{L^p(\Omega)}$$
(2) For any $\Omega' \subset\subset \Omega$ open and nonempty there exists $C = C(\Omega, \Omega')$ such that for all $u \in W^{1,p}(\Omega)$ we have
$$\|u - (u)_{\Omega'}\|_{L^p(\Omega)} \leq C(\Omega, \Omega')\|\nabla u\|_{L^p(\Omega)}$$
(3) There exists $C = C(\Omega)$ such that for all $u \in W_0^{1,p}(\Omega)$
$$\|u\|_{L^p(\Omega)} \leq C(\Omega)\|\nabla u\|_{L^p(\Omega)}$$

If $\Omega = B(x, r)$ (and in the second claim $\Omega' = B(x, \lambda r)$) then $C(\Omega) = C(B(0, 1))\, r$ (and for the second claim: $C(\Omega, \Omega') = C(B(0, 1), B(0, \lambda))\, r$).

Proof. The last claim can be proven by a scaling argument, and it is given as an exercise, Exercise 13.40.

Regarding the first claim, we simply let
$$K := \{u \in W^{1,p}(\Omega), (u)_\Omega = 0\}.$$
By Rellich's theorem, Theorem 13.35 this is a closed cone in $W^{1,p}$. Observe that if $u \in K$ is constant, $u \equiv C$ then $(u)_\Omega = C = 0$ by assumption, so $C = 0$. That is, the only constant function in K is the zero-function. Clearly $u - (u)_\Omega$ belongs to K, so we get the claim.

Regarding the second claim, we argue similarly setting
$$K := \{u \in W^{1,p}(\Omega), (u)_{\Omega'} = 0\}.$$

Regarding the third claim, observe that $W_0^{1,p}(\Omega)$ is (by definition) a closed set, and since it is a linear space it is in particular a cone. Now if $u \in W_0^{1,p}(\Omega)$ is constant, $u \equiv c$ then u is in particular continuous, but then by the zero trace theorem, Theorem 13.32 $c \equiv 0$. Again, the only constant function in K is the zero-function. $\square$

Exercise 13.39 (Bramble-Hilbert – higher order Poincaré)**.** *Prove that for any $k \in \mathbb{N}$ and any smoothly bounded domain Ω we have the following:*

For each $u \in W^{k,p}(\Omega)$ there exists polynomial p of degree at most $k - 1$ such that
$$\|u - p\|_{L^p(\Omega)} \leq C\, \|\nabla^k u\|_{L^p(\Omega)}$$

- *Show this is a version Poincaré's inequality, Corollary 13.38, if $k = 1$.*
- *Prove the statement for $k = 2$ by choosing explicitely the polynomial*
- *Prove the statement by blow-up.*

Let us also illustrate how one can obtain a more precise dependency on the constants using a *scaling argument*

Exercise 13.40. *Denote by $B(x_0, r) \subset \mathbb{R}$ the ball centered at x and with radius r. Show that there exists a uniform constant $C = C(n,p)$ such that for any $B(x_0,r)$ and any $u \in W^{1,p}(B(x_0,r))$ we have*

$$\|u - (u)_{B(x_0,r)}\|_{L^p(B(x_0,r)} \leq C(n,p)\, r\, \|\nabla u\|_{L^p(B(x_0,r))}.$$

Hint: Prove the inequality for $B(0,1)$. To get it for general $u \in W^{1,p}(B(x,r))$ apply the $B(0,1)$ inequality to $v(x) := u((x - x_0)/r)$ (see also Lemma 13.43).

We have seen in Theorem 13.35 and used in the Poincaré inequality that $W^{1,p}_{loc}(\Omega)$ embedds *compactly* into $L^p_{loc}(\Omega)$. There is a meta-theorem/feeling that "above" any compact embedding there is a merely continuous embedding, for more precise versions of this effect see [Hajłasz and Liu, 2010].

In our case it is that $W^{1,p}$ embedds into L^{p^*} where p^* follows the following rule

$$1 - \frac{n}{p} = 0 - \frac{n}{p^*} \tag{13.20}$$

(we will see this numerology appear later again for Morrey and Sobolev-Poincaré embedding, Corollary 13.45 and Theorem 13.49). Observe that $p^* = \frac{np}{n-p} \in (1,\infty)$ for $p < n$. We set $p^* := \infty$ for $p \geq n$. p^* is called the *Sobolev exponent*. What happens if $p^* > n$ (which should be interpreted from this numerological point of view as $p^* > \infty$? Theorem 13.49 will tell us: u is Hölder continuous.

Theorem 13.41 (Sobolev inequality)**.** *Let $p \in [1,\infty)$ such that $p^* := \frac{np}{n-p} \in (1,\infty)$ (equivalently: $p \in [1,n)$). Then $W^{1,p}(\mathbb{R}^n)$ embedds into $L^{p^*}(\mathbb{R}^n)$. That is, if $u \in W^{1,p}(\mathbb{R}^n)$ then $u \in L^{p^*}(\mathbb{R}^n)$ and we have*[36]

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C(p,n)\ \|Du\|_{L^p(\mathbb{R}^n)}.$$

Exercise 13.42. *Take $u \equiv 1$ in $\mathbb{R}^n$. Show that*

(1) $u \notin L^q(\mathbb{R}^n)$ for any $q \in [1,\infty)$.
(2) $Du = 0$ (in distributional sense)
(3) Conclude that for $p \in [1,n)$ we don't have

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C\ \|Du\|_{L^p(\mathbb{R}^n)}.$$

(4) Why does the latter fact not contradict Theorem 13.41?

[36]The optimal constant $C(p,n)$ has actually a geometric meaning, and is related to the isoperimetric inequality, cf. [Talenti, 1976]

Proof of Theorem 13.41. There are more than one way to prove Sobolev's inequality. One (the "Harmonic Analysis" one) is by convolution, using the Riesz potential representation and boundedness of Riesz transform on L^p-spaces. It is very strong and general but beyond the scope of these lectures.

The one we present here is an elegant trick due to Nirenberg (here we are again!). It is much less stable, relies heavily on the structure of $\mathbb{R}^n$, etc., but it obtains the case $p = 1$ (that in general is much more difficult to obtain), see e.g. [Schikorra et al., 2017].

By approximation it suffices to assume that $u \in C_c^\infty(\mathbb{R}^n)$.

Let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Then we have by the fundamental theorem of calculus

$$u(x_1, x_2, \ldots, x_n) = u(y_1, x_2, \ldots, x_n) + \int_{x_1}^{y_1} \partial_1 u(z_1, x_2, \ldots, x_n)\, dz_1.$$

Taking y_1 large enough we have $u(y_1, x_2, \ldots, x_n) = 0$, since $\operatorname{supp} u \subset\subset \mathbb{R}^n$. Thus we obtain the estimate

$$|u(x_1, x_2, \ldots, x_n)| \leq \int_{\mathbb{R}} |Du(z_1, x_2, \ldots, x_n)|\, dz_1.$$

The same way we obtain, for any $\ell = 1, \ldots, n$,

$$|u(x_1, x_2, \ldots, x_n)| \leq \int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell.$$

Multiplying these estimates for $\ell = 1, \ldots, n$ we obtain

$$|u(x_1, x_2, \ldots, x_n)|^n \leq \Pi_{\ell=1}^n \int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell.$$

Now we prove the case $p = 1$, when $p^* = \frac{n}{n-1}$. We have

$$\int_{\mathbb{R}} |u(x_1, x_2, \ldots, x_n)|^{\frac{n}{n-1}} dx_1 \leq \int_{\mathbb{R}} \Pi_{\ell=1}^n \left(\int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell \right)^{\frac{1}{n-1}} dx_1$$

$$\leq \left(\int_{z_1 \in \mathbb{R}} |Du(z_1, x_2, \ldots, x_n)|\, dz_1 \right)^{\frac{1}{n-1}} \int_{\mathbb{R}} \Pi_{\ell=2}^n \left(\int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell \right)^{\frac{1}{n-1}} dx_1$$

Now by Hölder's inequality[37]

$$\int_{\mathbb{R}} \Pi_{\ell=2}^n \left(\int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell \right)^{\frac{1}{n-1}} dx_1 \leq \left(\Pi_{\ell=2}^n \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell\, dx_1 \right)^{\frac{1}{n-1}}$$

and thus

$$\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 \leq \left(\int_{z_1 \in \mathbb{R}} |Du(z_1, x_2, \ldots, x_n)|\, dz_1 \right)^{\frac{1}{n-1}} \left(\Pi_{\ell=2}^n \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \ldots, z_\ell, \ldots x_n)|\, dz_\ell\, dx_1 \right)^{\frac{1}{n-1}}$$

[37] the generalized version for $k := n-1$ and all $p_i := n-1$: whenever $p_1, \ldots, p_k \in [1, \infty]$ and $\sum_i \frac{1}{p_i} = 1$,

$$\int_{\mathbb{R}^d} \Pi_{i=1}^k |f_i| \leq \Pi_{i=1}^k \left(\int_{\mathbb{R}^d} |f_i|^{p_i} \right)^{\frac{1}{p_i}}$$

Now we integrate this with respect to x_2, and again by Hölder's inequality,

$$
\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 dx_2 \\
\le &\int_{x_2\in\mathbb{R}} \left(\int_{z_1\in\mathbb{R}} |Du(z_1,x_2,\ldots,x_n)|\, dz_1\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=2}^n \int_{\mathbb{R}}\int_{\mathbb{R}} |Du(x_1,x_2,\ldots,z_\ell,\ldots x_n)|\, dz_\ell\, dx_1\right)^{\frac{1}{n-1}} \ dx_2 \\
\le &\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du(x_1,z_2,x_3,\ldots x_n)|\, dz_2\, dx_1\right)^{\frac{1}{n-1}} \\
&\quad\cdot \int_{x_2\in\mathbb{R}} \left(\int_{z_1\in\mathbb{R}} |Du(z_1,x_2,\ldots,x_n)|\, dz_1\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=3}^n \int_{\mathbb{R}}\int_{\mathbb{R}} |Du(x_1,x_2,\ldots,z_\ell,\ldots x_n)|\, dz_\ell\, dx_1\right)^{\frac{1}{n-1}} \ dx_2 \\
\le &\left(\int_{\mathbb{R}}\int_{\mathbb{R}} |Du(x_1,z_2,x_3,\ldots x_n)|\, dz_2\, dx_1\right)^{\frac{1}{n-1}} \\
&\quad\cdot \left(\int_{x_2\in\mathbb{R}}\int_{z_1\in\mathbb{R}} |Du(z_1,x_2,\ldots,x_n)|\, dz_1\, dx_2\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=3}^n \int_{x_2\in\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}} |Du(x_1,x_2,\ldots,z_\ell,\ldots x_n)|\, dz_\ell\, dx_1, dx_2\right)
\end{aligned}
$$

If $n = 2$ we are done (the $\Pi_{\ell=3}^n$-term is one). If $n \ge 3$ we see a pattern, continuing to integrate in $x_3,\ldots,x_n$ we obtain

$$
\begin{aligned}
\int_{\mathbb{R}}\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 dx_2 \le &\Pi_{\ell=1}^n \left(\int_{\mathbb{R}^n} |Du(x_1,\ldots,x_{\ell-1},z_\ell,x_{\ell+1}\ldots x_n)| dx_1,\ldots,x_{\ell-1},z_\ell,x_{\ell+1}\ldots x_n)\right)^{\frac{1}{n-1}} \\
= &\left(\int_{\mathbb{R}^n} |Du|\right)^{\frac{n}{n-1}}.
\end{aligned}
$$

Taking the exponent $\frac{n-1}{n}$ on both sides we obtain

$$
\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \le \|Du\|_{L^1(\mathbb{R}^n)} \tag{13.21}
$$

This is the claim for $p = 1$ (i.e. $p^* = \frac{n}{n-1}$).

The general claim follows when we apply the $p = 1$ Sobolev inequality to $v := |u|^\gamma$ for some $\gamma > 1$ that we choose later. We have

$$
|Dv| = |D|u|^\gamma| \le \gamma |u|^{\gamma-1}|Du|,
$$

thus (13.21) applied to v

$$
\||u|^\gamma\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \le C(\gamma)\||u|^{\gamma-1}\,|Du|\|_{L^1(\mathbb{R}^n)} \tag{13.22}
$$

Now observe that

$$
\||u|^\gamma\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} = \|u\|^\gamma_{L^{\gamma\frac{n}{n-1}}(\mathbb{R}^n)}
$$

Moreover, by Hölder's inequality, $p' = \frac{p}{p-1}$,

$$
\||u|^{\gamma-1}\,|Du|\|_{L^1(\mathbb{R}^n)} \le \||u|^{\gamma-1}\|_{L^{p'}(\mathbb{R}^n)}\|Du\|_{L^p(\mathbb{R}^n)} = \|u\|^{\gamma-1}_{L^{p'(\gamma-1)}(\mathbb{R}^n)}\|Du\|_{L^p(\mathbb{R}^n)}
$$

So (13.22) becomes

$$
\|u\|^\gamma_{L^{\gamma\frac{n}{n-1}}(\mathbb{R}^n)}\|u\|^{1-\gamma}_{L^{p'(\gamma-1)}(\mathbb{R}^n)} \le C(\gamma)\|Du\|_{L^p(\mathbb{R}^n)}
$$

Choosing $\gamma := \frac{p(n-1)}{n-p} > 1$ we have $\gamma\frac{n}{n-1} = p'(\gamma - 1) = p^*$, and then

$$\|u\|_{L^{p^*}(\mathbb{R}^n)} \leq C(\gamma)\|Du\|_{L^p(\mathbb{R}^n)}.$$

□

The relation between p and p^* in Theorem 13.41 is sharp in the following sense

Lemma 13.43. *Assume that $p, q \in (1,\infty)$ are such that for all $u \in C_c^\infty(\mathbb{R}^n)$*

$$\|u\|_{L^q(\mathbb{R}^n)} \leq C(p,n)\ \|Du\|_{L^p(\mathbb{R}^n)}. \tag{13.23}$$

Then $p = q^*$.

Proof. This is proven by a *scaling argument*. Assume (13.23) holds. Take an arbitrary $u \in C_c^\infty(\mathbb{R}^n)$ such that $\|Du\|_{L^p(\mathbb{R}^n)} \geq 1$, $\|u\|_{L^p(\mathbb{R}^n)} \geq 1$.

We rescale u and set for $\lambda > 0$,

$$u_\lambda(x) := u(\lambda x).$$

We apply (13.23) to u_λ. Observe that by substitution

$$\|u_\lambda\|_{L^q(\mathbb{R}^n)} = \lambda^{-\frac{n}{q}}\|u\|_{L^q(\mathbb{R}^n)},$$

and since $\nabla u_\lambda = \lambda(\nabla u)_\lambda$ we have

$$\|\nabla u_\lambda\|_{L^p(\mathbb{R}^n)} = \lambda^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

From (13.23) applied to u_λ we then obtain for any $\lambda > 0$,

$$\lambda^{-\frac{n}{q}}\|u\|_{L^q(\mathbb{R}^n)} \leq \lambda^{1-\frac{n}{p}}\|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Equivalently, setting $\Lambda := \|\nabla u\|_{L^q(\mathbb{R}^n)}/\|u\|_{L^q(\mathbb{R}^n)} > 0$ we obtain

$$\lambda^{0-\frac{n}{q}-(1-\frac{n}{p})} \leq \Lambda \quad \forall \lambda > 0$$

The exponent above the λ is exactly the numerology of (13.20)! In particular, if $q \neq p^*$ then $\sigma := 0 - \frac{n}{q} - (1 - \frac{n}{p}) \neq 0$, and we have

$$\lambda^\sigma \leq \Lambda \quad \forall \lambda > 0$$

If $\sigma > 0$ we let $\lambda \to \infty$, if $\sigma < 0$ we let $\lambda \to 0^+$ to get a contradiction. Thus, necessarily $\sigma = 0$, that is $q = p^*$. □

Corollary 13.44 (Sobolev-Poincaré embedding)**.** *Let $f \in W^{1,p}(\mathbb{R}^n)$, $1 \leq p < n$. For any $q \in [p, p^*]$ we have $f \in L^q(\mathbb{R}^n)$ with the estimate*

$$\|f\|_{L^q(\mathbb{R}^n)} \leq C(q,n)\ \Big(\|f\|_{L^p(\mathbb{R}^n)} + \|Df\|_{L^p(\mathbb{R}^n)}\Big).$$

Proof. We first claim that

$$\|f\|^q_{L^q(\mathbb{R}^n)} \leq C(q,n,p) \left(\|f\|^p_{L^p(\mathbb{R}^N)} + \|Df\|^{p^*}_{L^p(\mathbb{R}^n)}\right). \tag{13.24}$$

Clearly, the claim holds for $q = p$ and, by Theorem 13.41, for $q = p^*$.

Now observe that for $q \in [p, p^*]$ we can estimate the L^q-norm by the L^p-norm and the L^{p^*}-norm (this technique is called *interpolation*).

$$\int_{\mathbb{R}^n} |f|^q = \int_{\mathbb{R}^n} |f|^q \chi_{|f|>1} + \int_{\mathbb{R}^n} |f|^q \chi_{|f|\leq 1} \leq \int_{\mathbb{R}^n} |f|^{p^*} + \int_{\mathbb{R}^n} |f|^p.$$

That is,

$$\|f\|^q_{L^q(\mathbb{R}^n)} \leq \|f\|^p_{L^p(\mathbb{R}^n)} + \|f\|^{p^*}_{L^{p^*}(\mathbb{R}^n)} \precsim \|f\|^p_{L^p(\mathbb{R}^n)} + \|Df\|^{p^*}_{L^p(\mathbb{R}^n)}.$$

This proves (13.24).

How do we get the main claim? Well from (13.24) we find

$$\sup_{f \in W^{1,p}(\mathbb{R}^n):\ \|f\|_{W^{1,p}(\mathbb{R}^n)}\leq 1} \|f\|_{L^q(\mathbb{R}^n)} \leq \tilde{C}(p,n,q).$$

So in particular we have for any $f \in W^{1,p}(\mathbb{R}^n)$.

$$\|f/\|f\|_{W^{1,p}(\mathbb{R}^n)}\|_{L^q(\mathbb{R}^n)} \leq \tilde{C}(p,n,q)$$

Thus,

$$\|f\|_{L^q(\mathbb{R}^n)} \leq \tilde{C}(p,n,q)\, \|f\|_{W^{1,p}(\mathbb{R}^n)}.$$

□

Corollary 13.45 (Sobolev-Poincaré embedding on domains)**.** *Let $\Omega \subset \mathbb{R}^n$ and $\partial\Omega$ be C^1 (if $n = 1$ assume that Ω is an interval). For $1 \leq p < n$ we have for any $u \in W^{1,p}(\Omega)$,*

$$\|u\|_{L^{p^*}(\Omega)} \leq C(\Omega) \left(\|u\|_{L^p(\Omega)} + \|Du\|_{L^p(\Omega)}\right)$$

Also, for any $q \in [p, p^*]$

$$\|u\|_{L^q(\Omega)} \leq C(\Omega, q)\, \|u\|_{W^{1,p}(\Omega)}.$$

If moreover $\Omega \subset\subset \mathbb{R}^n$ and $u \in W^{1,p}_0(\Omega)$ then

$$\|u\|_{L^{p^*}(\Omega)} \leq C(\Omega)\, \|Du\|_{L^p(\Omega)}.$$

Lastly, if $1 \leq p < \infty$ and $\Omega \subset\subset \mathbb{R}^n$, $u \in W^{1,p}(\Omega)$ then for any $q \in [1, p^*]$ (if $p < n$) or for any $q \in [1, \infty)$ (if $p \geq n$)

$$\|u\|_{L^q(\Omega)} \leq C(\Omega, q, p, n)\, \|u\|_{W^{1,p}(\Omega)}.$$

Proof. By the extension theorem, Theorem 13.27, we can extend u to $Eu \in W^{1,p}(\mathbb{R}^n)$. Then from Sobolev inequality, Theorem 13.41, we get

$$\|u\|_{L^{p^*}(\Omega)} \leq \|Eu\|_{L^{p^*}(\mathbb{R}^n)} \precsim \|DEu\|_{L^p(\mathbb{R}^n)} \precsim C(\Omega)\|u\|_{W^{1,p}(\Omega)} \leq C(\Omega) \left(\|u\|_{L^p(\Omega)} + \|Du\|_{L^p(\Omega)}\right).$$

The second claim follows from the same argument using Corollary 13.44. Indeed from that we obtain For any $\Lambda > 0$ there exists a constant $C(\Omega, q, \Lambda)$ such that

$$\|u\|_{L^q(\Omega)} \leq C(\Omega, q, \Lambda) \quad \forall u: \ \|u\|_{W^{1,p}(\Omega)} \leq \Lambda.$$

Setting $\Lambda = 1$ and applying this inequality to $u/\|u\|_{W^{1,p}(\Omega)}$ we conclude.

The third claim follows from Poincaré inequality, Corollary 13.38, since for $u \in W_0^{1,p}(\Omega)$, $\Omega \subset\subset \mathbb{R}^n$ we have

$$\|u\|_{L^p(\Omega)} \leq C(\Omega)\|Du\|_{L^p(\Omega)}.$$

The last claim follows by additionall using Hölder's inequality: if $p < n$, $q \in [1, p^*]$,

$$\|u\|_{L^q(\Omega)} \leq C(|\Omega|, q, p)\|u\|_{L^{p^*}(\Omega)} \leq C(|\Omega|, q, p)\|u\|_{W^{1,p}(\Omega)}.$$

If $p \geq n$ and $n \geq 2$, and $q \in [p, \infty)$ we can find $1 \leq r \leq p$ such that $\infty > r^* = \frac{nr}{n-r} > q$. Thus, from first Hölder's inequality, then Sobolev inequality, and then again Hölder's inequality,

$$\|u\|_{L^q(\Omega)} \leq C(|\Omega|, q, p)\|u\|_{L^{r^*}(\Omega)} \leq C(|\Omega|, q, p)\|u\|_{W^{1,r}(\Omega)} \leq C(|\Omega|, q, p)\|u\|_{W^{1,p}(\Omega)}.$$

So the only remaining case is $n = 1$. Then by assumption $\Omega = (a, b)$. From the fundamental theorem of calulus

$$|f(x) - f(y)| \leq \int_{(a,b)} |f'(z)|dz \quad \forall x, y \in (a, b).$$

Thus

$$|f(x) - |b-a|^{-1} \int_{(a,b)} f(y)| \leq \int_{(a,b)} |f'(z)|dz \quad \forall x \in (a, b).$$

Thus,

$$|f(x)| \leq |b-a|^{-1}\|f\|_{L^1(a,b)} + \|f'\|_{L^1(a,b)} \quad \forall x \in (a, b)$$

That is

$$\|f\|_{L^\infty((a,b))} \leq C(a, b)\|f\|_{W^{1,1}((a,b))}.$$

In particular, by Hölder's inequality, for any $q, p \in [1, \infty]$.

$$\|f\|_{L^q((a,b))} \leq C(a, b, p, q)\|f\|_{W^{1,p}((a,b))}.$$

$\square$

Theorem 13.46 (Sobolev Embedding). *Let $\Omega \subset\subset \mathbb{R}^n$ be open, $\partial\Omega \in C^{0,1}$, $k \geq \ell$ for $k, \ell \in \mathbb{N} \cup \{0\}$, and $1 \leq p, q < \infty$, such that (compare with (13.20))*

$$k - \frac{n}{p} \geq \ell - \frac{n}{q}. \tag{13.25}$$

Then the identity is a continuous embedding $W^{k,p}(\Omega) \hookrightarrow W^{\ell,q}(\Omega)$. That is,

$$\|u\|_{W^{\ell,q}(\Omega)} \lesssim C(k, \ell, p, q)\, \|u\|_{W^{k,p}(\Omega)} \tag{13.26}$$

If $k > \ell$ and we have the strict inequality

$$k - \frac{n}{p} > \ell - \frac{n}{q}, \tag{13.27}$$

then the embedding above is compact. That is, whenever $(u_i)_{i\in\mathbb{N}} \subset W^{k,p}(\Omega)$ *such that*

$$\sup_i \|u_i\|_{W^{k,p}(\Omega)} < \infty$$

then there exists a subsequence $(u_{i_j})_{j\in\mathbb{N}}$ *such that* $(u_{i_j})_{j\in\mathbb{N}}$ *is convergent in* $W^{\ell,q}(\Omega)$*.*

Proof. If $k = \ell$, then (13.25) implies $p \geq q$. Thus, in that case (13.26) follows from the Hölder's inequality:

$$\|u\|_{W^{\ell,q}(\Omega)} \leq C(|\Omega|, n)\ \|u\|_{W^{\ell,p}(\Omega)} = C(|\Omega|, n)\ \|u\|_{W^{k,p}(\Omega)}$$

Next we us assume $k = \ell + 1$. Then (13.25) implies that $q \leq p^*$ (if $p < n$) or $q < \infty$ (for $p > n$), wher we recall the Sobolev exponent $p^* := \frac{np}{n-p}$. Then by Sobolev inequality, Corollary 13.45,

$$\|f\|_{L^q(\Omega)} \leq C(q, \Omega)\|f\|_{W^{1,p}(\Omega)}.$$

Applying this inequality to $f := \partial^\gamma u$ for $|\gamma| \leq \ell$ we obtain (13.26) for $k = \ell + 1$, namely for $q \leq p^*$,

$$\|u\|_{W^{\ell,q}(\Omega)} \leq C(p, q, \Omega)\|u\|_{W^{\ell+1,p}(\Omega)}$$

More generally If $k = \ell + N$ for some $N \in \mathbb{N}$, set $r_i := (r_{i-1})^*$ for $i = 1, \ldots, N$ with $r_0 := p$. This works well if all of the $r_i^* \neq \infty$ (otherwise we choose $r_i \leq (r_{i_1})^*$ and $r_0 < p$, but large enough such that $r_N > q$). Then (13.25) implies that $q \leq r_N$, and we get first by Hölder's inequality then by the argument above iterated

$$\|u\|_{W^{\ell,q}(\Omega)} \precsim \|u\|_{W^{\ell,r_N}(\Omega)} \precsim \|u\|_{W^{\ell+1,r_{N-1}}(\Omega)} \precsim \ldots \precsim \|u\|_{W^{k,r_0}(\Omega)} \precsim \|u\|_{W^{k,p}(\Omega)}.$$

This proves the continuous embedding, (13.27) in full generality.

As for the compact embedding, it suffices to assume $k = \ell+1$. This is because combinations of continuous and compact embeddings are compact, so if we show the compactness of the embedding satisfying (13.27) for $k = \ell + 1$ then we can build a chain of embeddings as above to get a compact embedding for all $k > \ell$.

Moreover, we can assume w.l.o.g. $k = 1$, $\ell = 0$. The general case then follows by considering $\partial^\gamma u$ for $|\gamma| \leq \ell$.

So let $1 \leq q < p^*$ (i.e. (13.27) and assume that we have a sequence (u_i) such that

$$\sup_i \|u_i\|_{W^{1,p}(\Omega)} < \infty.$$

Fix $r \in (q, p^*)$ (if $p \geq n$ then $r > q$. By Sobolev's inequality, Corollary 13.45,

$$\Lambda := \sup_i \|u_i\|_{L^r(\Omega)} < \infty. \tag{13.28}$$

By Rellichs theorem, Theorem 13.35, we can find a subsequence u_{i_j} that is strongly convergent in $L^p(\Omega)$ and in particular we can choose the subsequence such that u_{i_j} converges pointwise a.e. to some $u \in L^q(\Omega)$ (that u belongs to L^r, and thus to L^q follows from the weak compactness, Theorem 12.13, or Fatou's lemma).

Now we use Vitali's convergence theorem, Theorem 3.59. To show the uniform absolute continuity of the integral let $\varepsilon > 0$ and for some δ to be chosen (independent of j) let $E \subset \Omega$ be measurable with $|E| < \delta$. Then we have by Hölder's inequality (recall Λ from (13.28))

$$\sup_j \|u_{i_j}\|_{L^q(E)} \leq |E|^{\frac{1}{q}-\frac{1}{r}} \sup_j \|u_{i_j}\|_{L^r(E)} \leq \delta^{\frac{1}{q}-\frac{1}{r}}\Lambda.$$

So if we choose $\delta = \delta(\varepsilon, \Lambda) > 0$ small, so that

$$\delta^{\frac{1}{q}-\frac{1}{p}}\Lambda < \varepsilon,$$

then

$$\sup_j \|u_{i_j}\|_{L^q(E)} < \varepsilon \quad \text{whenever } E \subset \Omega \text{ measurable and } |E| < \delta.$$

This is uniform absolute continuity, and by Vitali's theorem u_{i_j} is convergent in $L^q(\Omega)$. This shows compactness, and Sobolev's embedding theorem is proven. □

Our next goal is Morrey's embedding theorem, Theorem 13.49. For this we use a characterization of Hölder functions by so-called *Campanato spaces*.

Theorem 13.47 (Campanato's theorem). *Let $u \in L^1(\mathbb{R}^n)$ and assume that for some $\lambda > 0$*

$$\Lambda := \sup_{B(x,r)\subset\mathbb{R}^n} r^{-\lambda} \fint_{B(x,r)} |u - (u)_{B(x,r)}| < \infty, \tag{13.29}$$

where

$$(u)_{B(x,r)} = \fint_{B(x,r)} u.$$

Then $u \in C^{\lambda}_{loc}(\mathbb{R}^n)$ and we have for some uniform constant $C = C(n,\lambda) > 0$

$$\sup_{x\neq y} \frac{|u(x)-u(y)|}{|x-y|^{\lambda}} \leq C\ \Lambda.$$

Remark 13.48. The converse also holds, if $u \in C^{\alpha}$ than $\Lambda < [u]_{C^{\alpha}}$, which is an easy exercise to check.

Proof of Theorem 13.47. First we claim that for any $R > 0$ and almost any $x \in \mathbb{R}^n$ we have for some uniform constant $C > 0$

$$|u(x) - (u)_{B(x,R)}| \leq C\ R^{\lambda}\Lambda. \tag{13.30}$$

To see this, observe that for almost every $x \in \mathbb{R}^n$, by Lebesgue's theorem, Theorem 5.18, $\lim_{k\to\infty}(u)_{B(x,2^{-k}R)} = u(x)$. Thus, by a telescoping sum

$$|u(x) - (u)_{B(x,R)}| \leq \sum_{k=0}^{\infty} \left|(u)_{B(x,2^{-k}R)} - (u)_{B(x,2^{-(k+1)}R)}\right| \tag{13.31}$$

Now,

$$\begin{aligned}
&\left|(u)_{B(x,2^{-k}R)} - (u)_{B(x,2^{-(k+1)}R)}\right| \\
\leq &\fint_{B(x,2^{-(k+1)}R)} \left|u(z) - (u)_{B(x,2^{-k}R)}\right| \\
\leq &\underbrace{\frac{|B(x,2^{-k}R)|}{|B(x,2^{-k+1}R)|}}_{=C(n)} \fint_{B(x,2^{-k}R)} \left|u(z) - (u)_{B(x,2^{-k}R)}\right| \\
\overset{(13.29)}{\leq} &C(n)\,\Lambda\,(2^{-k}R)^{\lambda}.
\end{aligned}$$

Plugging this into (13.31) we get

$$|u(x) - (u)_{B(x,R)}| \leq C(n)\Lambda R^{\lambda} \sum_{k=0}^{\infty} 2^{-k\lambda} \overset{\lambda \geq 0}{=} C(\lambda,n)\Lambda\, R^{\lambda},$$

i.e. (13.30) is established.

Now let $x, y \in \mathbb{R}^n$. Set $R := |x-y|$. Then

$$\begin{aligned}
|u(x) - u(y)| \leq &|u(x) - (u)_{B(x,R)}| + |u(x) - (u)_{B(y,R)}| + |(u)_{B(y,R)} - (u)_{B(x,R)}| \\
\overset{(13.30)}{\leq} &C(n,\lambda)|x-y|^{\lambda} + |(u)_{B(y,R)} - (u)_{B(x,R)}|.
\end{aligned} \tag{13.32}$$

We have to estimate the last term, which we do as above: Observe that $B(x,2R) \supset B(y,R) \cup B(x,R)$,

$$\begin{aligned}
&|(u)_{B(y,R)} - (u)_{B(x,R)}| \\
\leq &\fint_{B(x,R)} \fint_{B(y,R)} |u(z_1) - u(z_2)| dz_1 dz_2 \\
\leq &\underbrace{\frac{|B(x,2R)|}{|B(y,R)|} \frac{|B(x,2R)|}{|B(x,R)|}}_{=C(n)} \fint_{B(x,2R)} \fint_{B(x,2R)} |u(z_1) - u(z_2)| dz_1 dz_2 \\
= &C(n) \fint_{B(x,2R)} \fint_{B(x,2R)} |u(z_1) - (u)_{B(x,2R)}| dz_1 dz_2 + C(n) \fint_{B(x,2R)} \fint_{B(x,2R)} |(u)_{B(x,2R)} - u(z_2)| dz_1 dz_2 \\
= &C(n) \fint_{B(x,2R)} \fint_{B(x,2R)} |u(z_1) - (u)_{B(x,2R)}| dz_1 dz_2 + C(n) \fint_{B(x,2R)} \fint_{B(x,2R)} |(u)_{B(x,2R)} - u(z_2)| dz_1 dz_2 \\
= &2C(n) \fint_{B(x,2R)} |u(\tilde{z}) - (u)_{B(x,2R)}| d\tilde{z} \\
\overset{(13.29)}{\leq} &2C(n,\lambda) R^{\lambda}.
\end{aligned}$$

Since $R = |x-y|$, together with (13.32) we have shown

$$|u(x) - u(y)| \leq C(n,\lambda)|x-y|^{\lambda},$$

and can conclude. $\square$

Theorem 13.49 (Morrey Embedding). *Let $\Omega \subset\subset \mathbb{R}^n$ with $\partial\Omega \in C^k$, $k \in \mathbb{N}$. Assume that for $p \in (1,\infty)$, $\alpha \in (0,1)$ and $\ell < k$ we have*
$$k - \frac{n}{p} \geq \ell + \alpha.$$
Then the embedding $W^{k,p}(\Omega) \hookrightarrow C^{\ell,\alpha}(\overline{\Omega})$ is continuous.

If $k - \frac{n}{p} > \ell + \alpha$ then the embedding is compact.

Proof. Let $u \in W^{k,p}(\Omega)$. By Extension Theorem, Theorem 13.27, we can assume $u \in W^{k,p}(\mathbb{R}^n)$ and $\operatorname{supp} u \subset\subset B(0,R)$ for some large $R > 0$.

As in the Sobolev theorem it suffices to assume $\ell = k-1$, and indeed we can reduce to the case $k = 1$ and $\ell = 0$.

We use Campanato's Theorem, Theorem 13.47. For $B(x,r) \subset \mathbb{R}^n$, we have by Poincaré's inequality, Corollary 13.38, and then Hölder's inequality,
$$\fint_{B(x,r)} |u-(u)_{B(x,r)}| \leq r^{1-\lambda} \fint_{B(x,r)} |Du| = Cr^{1-n}\int_{B(x,r)} |Du| \leq Cr^{1-n} r^{n-\frac{n}{p}} \left(\int_{B(x,r)} |Du|^p\right)^{\frac{1}{p}}$$
That is,
$$\sup_{B(x,r)\subset\mathbb{R}^n} r^{-\left(1-\frac{n}{p}\right)} \fint_{B(x,r)} |u-(u)_{B(x,r)}| \leq C\|Du\|_{L^p(\mathbb{R}^n)}$$
Thus, by Campanato's theorem, if $1 - \frac{n}{p} = 0 + \alpha \in (0,1)$, then (using also the extension theorem estimate),
$$[u]_{C^\alpha(\Omega)} \leq [u]_{C^\alpha(\mathbb{R}^n)} \leq C\|u\|_{W^{1,p}(\Omega)},$$
which is the continuity of the embedding of $W^{1,p}(\Omega)$ in $C^{0,\alpha}$ if $1 - \frac{n}{p} = 0 + \alpha$.

If on the other hand $1 - \frac{n}{p} > 0 + \alpha$, then we use Arzela-Ascoli to show that the embedding $L^\infty \cap C^\beta(\mathbb{R}^n) \hookrightarrow L^\infty \cap C^\alpha(\mathbb{R}^n)$ is compact if $\beta > \alpha$, and from this we conclude the compactness of the embedding $W^{1,p}(\Omega) \hookrightarrow C^\alpha(\overline{\Omega})$ if $1 - \frac{n}{p} > \alpha$. □

In general $W^{1,n}$-functions in $\mathbb{R}^n$ may not be continuous, if $n \geq 2$, $\log\log|x|$ is the example to have in mind. However recall that for $n = 1$ we have

Proposition 13.50. *Let $f \in W^{1,1}((a,b))$ then $f \in C^0(a,b)$ (in the sense of representative).*

Proof. Approximate f by $f_k \in C^\infty([a_1,b_1])$ where $a < a_1 < b_1 < b$ is taken arbitary.

We have by the fundamental theorem of Calculus for all $x,y \in [a_1,b_1]$,
$$f_k(x) - f_k(y) = \int_x^y f_k'(z)dz.$$
That is
$$|f_k(x) - f_k(y)| \leq \int_{(a_1,b_1)} |f_k'(z) - f'(z)|dz + \int_x^y |f'(z)|dz.$$

Now let $\varepsilon > 0$. By absolute continuity of the integral and since $f' \in L^1$, there exists a $\delta_1 > 0$ such that

$$\int_x^y |f'(z)|dz < \frac{\varepsilon}{4} \quad \forall x, y \in [a_1, b_1], |x - y| < \delta_1.$$

By L^1-convergence $f_k' \to f'$ there exists some $K \in \mathbb{N}$ such that

$$\sup_{k \geq K} \int_{(a_1,b_1)} |f_k'(z) - f'(z)|dz < \frac{\varepsilon}{4}.$$

Lastly let $\delta_2 < \delta_1$ such that for all (finitely many!) $k \in \{1, \ldots, K\}$ we have

$$|f_k(x) - f_k(y)| < \frac{\varepsilon}{4} \quad \forall |x - y| < \delta_2 \quad k \in \{1, \ldots, K\}.$$

Then we have

$$|f_k(x) - f_k(y)| < \varepsilon \quad \forall k \in \mathbb{N}, \quad |x - y| < \delta_2.$$

That is, $(f_k)_{k\in\mathbb{N}}$ is equicontinuous on $[a_1, b_1]$.

Since $f_k(x)$ converges (up to going to a subsequence which we will not relabel) a.e. to $f(x)$, we can pick some $x \in [a_1, b_1]$ for which $f(x) < \infty$ and then conclude that by equicontinuity (up to taking again a subsequence not relabeled)

$$\sup_k |f_k(y) - f(x)| \leq \sup_k (|f_k(x) - f_k(y)| + |f(x)|) \leq C < \infty.$$

That is (f_k) satisfies the conditions for Arzela-Ascoli, thus it converges uniformly to a continuous function g.

Since f_k also converges a.e. to f we conclude that $g = f$ a.e. – i.e. f as a continuous representative. □

One can also show that $W^{n,1}$-maps are continuous in $\mathbb{R}^n$.

13.7. **Fun inequalities: Ehrling's lemma, Gagliardo-Nirenberg inequality, Hardy's inequality.**

We know Sobolev and Poincaré inequality in various versions. There are many more cool inequalities.

13.7.1. *Ehrling's Lemma.*

Theorem 13.51 (Functional Analytic Ehrling's lemma). *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ be three Banach spaces which are subspaces of each other $X \subset Y \subset Z$ with the following properties.*

- *X is compactly embedded in Y, that is $X \subset Y$ and every $\|\cdot\|_X$-bounded sequence $(x_k)_k \subset X$, $\sup_k \|x_k\|_X < \infty$, has a strongly $\|\cdot\|_Y$-convergent subsequence $(x_{k_i})_{i\in\mathbb{N}}$, i.e. for some $y \in Y$ and $\|x_{k_i} - y\|_Y \xrightarrow{i\to\infty} 0$.*
- *Y is continuously embedded in Z, that is $Y \subset Z$ and there exists $\Lambda > 0$ such that $\|y\|_Z \leq \Lambda \|y\|_Y$ for all $y \in Y$.*

Then for every $\varepsilon > 0$ there exists a constant $C(\varepsilon) > 0$ such that the following holds

$$\|x\|_Y \leq \varepsilon \|x\|_X + C(\varepsilon)\|x\|_Z \quad \forall x \in X.$$

Exercise 13.52. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be Banach spaces. Show that if $X \subset Y$ is compactly embedded, then $X \subset Y$ is continuously embedded.*

Proof of Theorem 13.51. This is once again a typical blow-up proof.

Fix $\varepsilon > 0$. Assume the claim is false, then for any $k \in \mathbb{N}$ there exists a "counterexample" $x_k \in X$ such that

$$\|x_k\|_Y > \varepsilon \|x_k\|_X + k\|x_k\|_Z \quad \forall k$$

Dividing this inequality by $\|x_k\|_Y$ (cannot be zero because of the strict inequality) and otherwise switching over to $\tilde{x}_k := \frac{x_k}{\|x_k\|_Y}$ we may assume w.l.o.g. $\|x_k\|_Y = 1$ for all k and thus

$$1 > \varepsilon \|x_k\|_X + k\|x_k\|_Z \quad \forall k.$$

In particular,

$$\sup_k \|x_k\|_X \leq \frac{1}{\varepsilon}$$

and

$$\lim_{k\to\infty} \|x_k\|_Z \leq \lim_{k\to\infty} \frac{1}{k} = 0.$$

By compactness $X \subset Y$ and since $(x_k)_k$ is $\|\cdot\|_X$-bounded, up to passing to a subsequence $(x_{k_i})_i$, we can assume w.l.o.g. that there exists $y \in Y$ such that $\|x_k - y\|_Y \xrightarrow{k\to\infty} 0$. In particular since $\|x_k\|_Y = 1$ we find that $\|y\|_Y = 1$. By the continuous embedding $Y \subset Z$ we also have $\|x_k - y\|_Z \xrightarrow{k\to\infty} 0$, but since $\|x_k\|_Z \xrightarrow{k\to\infty} 0$ we have $y = 0$, which contradicts $\|y\|_Y = 1$.

So there must have been some $k \in \mathbb{N}$ for which there was no counterexample x_k – and thus the claim is proven. $\square$

Corollary 13.53 (Sobolev spaces version)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary, and $0 \leq \ell < k$, $p \in [1, \infty)$. Then for any $\varepsilon > 0$ there exists $C(\Omega, \varepsilon, p) > 0$ such that*

$$\|D^\ell u\|_{L^p(\Omega)} \leq \varepsilon \|D^k u\|_{L^p(\Omega)} + C(\Omega, \varepsilon, p)\|u\|_{L^p(\Omega)}$$

Proof. It suffices to prove this claim for $k = \ell + 1$. Indeed, then (we can always assume that ε is small!)

$$\begin{aligned}
\|D^\ell u\|_{L^p(\Omega)} \leq& \varepsilon \|D^{\ell+1} u\|_{L^p(\Omega)} + C_(\Omega, \varepsilon, p, \ell)\|u\|_{L^p(\Omega)} \\
\leq& \varepsilon^2 \|D^{\ell+2} u\|_{L^p(\Omega)} + (C(\Omega, \varepsilon, p, \ell) + \varepsilon C(\Omega, \varepsilon, p, \ell + 1))\|u\|_{L^p(\Omega)} \\
\leq& \ldots \\
\leq& \varepsilon^{k-\ell} \|D^{\ell+2} u\|_{L^p(\Omega)} + C(\Omega, \varepsilon, p, \ell, k)\|u\|_{L^p(\Omega)}.
\end{aligned}$$

So assume $k = \ell + 1$. Set

$$X := W^{\ell+1,p}(\Omega), \quad Y := W^{\ell,p}(\Omega), \quad Z = L^p(\Omega).$$

In view of Sobolev and Morrey's embedding, Theorem 13.46 and Theorem 13.49, all the conditions of Theorem 13.51 are met, so we have for any $\varepsilon > 0$ some $C(\varepsilon) > 0$ such that

$$\sum_{|\alpha| \leq \ell} \|D^\alpha u\|_{L^p} \leq \varepsilon \sum_{|\alpha| \leq \ell+1} \|D^\alpha u\|_{L^p} + C(\varepsilon)\|u\|_{L^p} \quad \forall u \in W^{\ell+1,p}(\Omega).$$

Now if $\varepsilon < \frac{1}{2}$ (which we can always assume), we can absorb the terms up to order ℓ on the right-hand side, namely

$$(1-\varepsilon) \sum_{|\alpha| \leq \ell} \|D^\alpha u\|_{L^p} \leq \varepsilon \sum_{|\alpha| = \ell+1} \|D^\alpha u\|_{L^p} + C(\varepsilon)\|u\|_{L^p} \quad \forall u \in W^{\ell+1,p}(\Omega).$$

thus

$$\sum_{|\alpha| \leq \ell} \|D^\alpha u\|_{L^p} \leq \frac{\varepsilon}{1-\varepsilon} \sum_{|\alpha| = \ell+1} \|D^\alpha u\|_{L^p} + \frac{1}{1-\varepsilon} C(\varepsilon)\|u\|_{L^p} \quad \forall u \in W^{\ell+1,p}(\Omega).$$

In particular, with only dimensional constant $\Lambda = \Lambda(n)$, for all $\varepsilon < \frac{1}{2}$,

$$\|D^\ell u\|_{L^p} \leq \Lambda\varepsilon \|D^{\ell+1} u\|_{L^p} + \tilde{C}(\varepsilon)\|u\|_{L^p} \quad \forall u \in W^{\ell+1,p}(\Omega).$$

This proves the claim. □

Exercise 13.54 (Equivalent norm for Sobolev spaces)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary, and $0 \leq \ell < k$, $p \in [1, \infty)$.*

Show that the two norms on $W^{k,p}(\Omega)$,

$$\|u\|_1 := \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^p}$$

and

$$\|u\|_2 := \|u\|_{L^p} + \|D^\alpha u\|_{L^p},$$

are equivalent.

13.7.2. *Gagliardo-Nirenberg inequality.* Very similar to Ehrling's lemma, Section 13.7.1, but with different p's (making compactness argument not work, because it is more of a Sobolev-type inequality) and its on $\mathbb{R}^n$ (though there exists also a domain version)

Theorem 13.55. *Let $1 < p, q, r < \infty$ and $\alpha \in (0, 1)$ such that*

$$\frac{1}{p} = \frac{j}{n} + \left(\frac{1}{r} - \frac{m}{n}\right)\alpha + \frac{1-\alpha}{q}.$$

and $\frac{j}{m} \leq \alpha \leq 1$.

Assume $u \in L^q(\mathbb{R}^n)$ and its distributional m-th derivative $D^m u \in L^r(\mathbb{R}^n)$. Then $D^j u \in L^p(\mathbb{R}^n)$ and for a constant $C = C(m, n, j, q, r, \alpha)$ we have

$$\|D^j u\|_{L^p(\mathbb{R}^n)} \leq C\, \|D^m u\|^{\alpha}_{L^r(\mathbb{R}^n)}\, \|u\|^{1-\alpha}_{L^q(\mathbb{R}^n)}.$$

FIGURE 13.1. Olga Ladyzhenskaya: 1922-2004 – Russian mathematician who obtained numerous breakthroughs in Partial Differential Equations, Fluid dynamics, Navier-Stokes equation. Without any doubt one of the most impressive mathematicians in Analysis in the 20th century!

There also exists versions in the limit cases $= \infty$, $= 1$, etc.

Proof. See e.g. [Leoni, 2017, Theorem 12.83.]. Here we just discuss that it is enough to show

$$\|D^j u\|_{L^p(\mathbb{R}^n)} \le C \left(\|D^m u\|_{L^r(\mathbb{R}^n)} + \|u\|_{L^q(\mathbb{R}^n)}\right)$$

Indeed, then we can obtain the claim by scaling: Apply the inequality to $u(\lambda\cdot)$, then we get

$$\lambda^{-\frac{n}{p}}\lambda^j\|D^j u\|_{L^p(\mathbb{R}^n)} \le C \left(\lambda^{-\frac{n}{r}}\lambda^m\|D^m u\|_{L^r(\mathbb{R}^n)} + \lambda^{-\frac{n}{q}}\|u\|_{L^q(\mathbb{R}^n)}\right)$$

That is

$$\|D^j u\|_{L^p(\mathbb{R}^n)} \le C \left(\lambda^{-\frac{n}{r}+\frac{n}{p}+m-j}\|D^m u\|_{L^r(\mathbb{R}^n)} + \lambda^{\frac{n}{p}-j-\frac{n}{q}}\|u\|_{L^q(\mathbb{R}^n)}\right)$$

So compute

$$\min_{\lambda\ge 0}\left(\lambda^{-\frac{n}{r}+\frac{n}{p}+m-j}\|D^m u\|_{L^r(\mathbb{R}^n)} + \lambda^{\frac{n}{p}-j-\frac{n}{q}}\|u\|_{L^q(\mathbb{R}^n)}\right).$$

(compute means: check $\lambda = 0$ and $\lambda = \infty$ tend to $+\infty$, then take the derivative in λ and check the critical case). □

13.7.3. *Ladyshenskaya inequality.* Ladyshenskaya's inequality is a special case of the Gagliardo-Nirenberg inequality Theorem 13.55, used often for the Navier-Stokes equations for the term $u \cdot \nabla u$:

$$\|u\|_{L^4(\mathbb{R}^2)} \lesssim \|u\|_{L^2(\mathbb{R}^2)}^{\frac{1}{2}} \|\nabla u\|_{L^2(\mathbb{R}^2)}^{\frac{1}{2}}.$$

and

$$\|u\|_{L^4(\mathbb{R}^3)} \lesssim \|u\|_{L^2(\mathbb{R}^3)}^{\frac{1}{4}} \|\nabla u\|_{L^2(\mathbb{R}^3)}^{\frac{3}{4}}.$$

the difference in powers is (to some extend) what makes Navier-Stokes equations in 3D more challenging than in 2D. There are appropriate versions on domains Ω, if the functions are zero on $\partial\Omega$.

13.7.4. *Hardy's inequality.* Hardy's inequality is somewhat a weighted Sobolev inequality.

Theorem 13.56. *Let $s \in (0,1)$, $p,q \in (1,\infty)$ such that*

$$\frac{1-s}{n} - \frac{n}{p} = \frac{n}{q}.$$

Then

$$\||x|^{-s} f(x)\|_{L^p(\mathbb{R}^n)} \leq C_{s,p} \|\nabla f\|_{L^q(\mathbb{R}^n)}.$$

Observe that for $s = 0$ we obtain the usual Sobolev inequality, Theorem 13.41.

See, e.g, [Evans, 2010, p.296, Theorem 7] or [Mironescu, 2018].

13.8. **Rademacher's theorem.** The following is *Rademacher's theorem* (usually stated for Lipschitz functions, cf. Theorem 13.24)

Theorem 13.57. *Let $p > n$. Let $u \in W^{1,p}(\Omega)$ for some open set $\Omega \subset \mathbb{R}^n$ and $1 \leq p \leq \infty$. We identify u with its continuous representative. Then for almost every $x \in \Omega$ u is differentiable, that is there exists $A = A(x) \in \mathbb{R}^n$ such that*

$$\lim_{y \to x} \frac{|u(y) - u(x) - A(y-x)|}{|y-x|} = 0$$

Moroever $A(x) = Du(x)$ (where $Du(x)$ denotes the distributional gradient which belongs to L^p) for almost every $x \in \Omega$.

We need the following estimate (which we could have used for Sobolev-Morrey embedding, Theorem 13.49).

Lemma 13.58 (Morrey's estimate)**.** *Let $p > n$. Then there exists a constant $C = C(n,p)$ such that the following holds: for any $v \in C^1(B(0,2r))$ we have for all $x, y \in B(x_0, r)$*

$$|v(y) - v(x)| \leq C\, r^{1-\frac{n}{p}} \left(\int_{B(x_0,2r)} |Dv(z)|^p \, dz \right)^{\frac{1}{p}}$$

Proof. We assume $r = 1$, the general case is part of Exercise 13.59. Arguing as in the proof of Theorem 13.47 we have (for every point, since v is continuous by Theorem 13.47)

$$\begin{aligned}|v(y) - v(x)| \leq C(n) \Big(&\sum_{k=0}^{\infty} \fint_{B(x,2^{-k})} |u(z) - (u)_{B(x,2^{-k})}| \\ &+ \sum_{k=0}^{\infty} \fint_{B(y,2^{-k})} |u(z) - (u)_{B(y,2^{-k})}| \\ &+ |(v)_{B(x,1)} - (v)_{B(y,1)}| \Big).\end{aligned}$$

Observe that

$$|(v)_{B(x,1)} - (v)_{B(y,1)}| \leq \fint_{B(0,2)} \fint_{B(0,2)} |v(z_1) - v(z_2)| \leq 2 \fint_{B(0,2)} \fint_{B(0,2)} |v(z_1) - (v)_{B(0,2)}|.$$

Now we use Poincaré inequality, Exercise 13.40 and have (constants change from equation to equation!)

$$\fint_{B(z,\rho)} |f - (f)_{B(z,\rho)}| \leq C\, \rho^{\frac{n}{p'} - n} \|f - (f)_{B(z,\rho)})\|_{L^p(B(z,\rho)} \leq C\, \rho^{1 + \frac{n}{p'} - n} \|\nabla f\|_{L^p(B(z,\rho))}$$

That is, we have

$$\fint_{B(y,2^{-k})} |u(z) - (u)_{B(y,2^{-k})}| \leq C 2^{-k\frac{p-n}{p}} \|\nabla u\|_{L^p(B(0,2))}.$$

Since $p > n$ we have that $\sum_{k=1}^{\infty} 2^{-k\frac{p-n}{p}} < \infty$ and can conclude. □

Exercise 13.59. *Use a scaling argument to finish the proof of Lemma 13.58 for general r.*

Proof of Theorem 13.57. Denote by $Du(x)$ the Lebesgue representative of $Du \in L^p(\Omega)$ we have from Lebesgue's differentiation theorem, Theorem 5.18 for almost every $x \in \Omega$

$$\lim_{r \to 0} \fint_{B(x,r)} |Du(x) - Du(z)|^p dz = 0.$$

Take x a point where this is true. Set (observe u is continuous by embedding theorem, since $p > n$)

$$v(y) := u(y) - u(x) - Du(x)(y - x).$$

Then $v \in W^{1,p}_{loc}(\Omega) \cap C^0(\Omega)$.

From Lemma 13.58, which by approximation holds for every y and r such that $y \in B(x, 4r) \subset \Omega$ (recall that x is fixed)

$$|v(y) - \underbrace{v(x)}_{=0}| \leq C\, r^{1-\frac{n}{p}} \left(\int_{B(x,2r)} |Dv(z)|^p dz \right)^{\frac{1}{p}} \leq C\, r^{1} \left(\fint_{B(x,2r)} |Dv(z)|^p dz \right)^{\frac{1}{p}}$$

So let $r = |x - y|$ then we have

$$\frac{|u(y) - u(x) - Du(x)(y - x)|}{|x - y|} \leq C \left(\fint_{B(x,2|x-y|)} |Dv(z)|^p dz \right)^{\frac{1}{p}}.$$

By assumption we have that

$$\left(\fint_{B(x,2|x-y|)} |Dv(z)|^p dz\right)^{\frac{1}{p}} \xrightarrow{|x-y|\to 0} 0,$$

so we have shown that u is differentiable in x.

$\square$

From Theorem 13.57 follows the usual Rademacher theorem.

Corollary 13.60 (Rademacher Theorem)**.** *If $f : \Omega \to \mathbb{R}$ is a Lipschitz map where Ω is open, then f is almost everywhere differentiable. Moreover pointwise and distributional derivative coincide a.e.*

Proof. From Theorem 13.24 we have that $f \in W^{1,\infty}_{loc}(\Omega)$ thus by Hölder's inequality $f \in W^{1,p}_{loc}(\Omega)$. Now this is a consequence of Theorem 13.57. $\square$

Remark 13.61. The set where Sobolev functions are differentiable can be made more precise than simply being "a.e". There is a notion of Sobolev-*capacity* of sets, and one can show that they are differentiable outside of a set of certain capacity zero.

14. Sobolev spaces between manifolds

14.1. **Short excursion on degree and Brouwer Fixed Point theorem.**

We denote by $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \quad |x| = 1\}$ the unit sphere in $\mathbb{R}^n$. By a slight abuse of notation we will call $\mathbb{B}^n = \{x \in \mathbb{R}^n : |x| \leq 1\}$ the *closed* unit ball in $\mathbb{R}^n$. Observe that $\mathbb{S}^{n-1} = \partial \mathbb{B}^n$.

Proposition 14.1. *There is no smooth map $\Phi : \overline{\mathbb{B}^n} \to \mathbb{S}^{n-1}$ such that*

$$\Phi(x) = x \quad \forall x \in \mathbb{S}^{n-1}.$$

Proof. The reason this is true is degree theory. The map $x : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ "winds around" the target exactly once. Now if we consider $\Phi_r\Big|_{\partial B(0,r)}$ then $r \mapsto \Phi_r$ transforms continuously the curve Φ_1 into the constant map Φ_0. However degree is a homotopy invariant, i.e. the "winding around" does not change under continuous changes – so we have a contradiction.

More precisely let $\varphi : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$. We define the degree (if $n = 2$ one can see this is the winding number)

$$\deg(\varphi) := \int_{B^n} \det(D\Phi),$$

where $\Phi : \mathbb{B}^n \to \mathbb{R}^n$ is any smooth extension of φ. We have to show that this is well-defined, i.e. for two different choices of Φ_1, Φ_2 both extensions of φ we have to show

$$\int_{B^n} \det(D\Phi_1) = \int_{B^n} \det(D\Phi_2).$$

Observe that the determinant is multilinear, so by subtracting the left from the right it actually suffices to show that

$$\int_{B^n} \det(D\psi^1|D\psi^2|\dots|D\psi^n) = 0$$

whenever *one* of the ψ^i satisfies $\psi^i = 0$ on $\partial\mathbb{B}^n$. By rearranging we can assume that $\psi^1 = 0$ on $\partial\mathbb{B}^n$. We will show this claim in two dimensions. In higher dimensions it gets more messy, but the principle stays the same: the co-factor of a gradient matrix is divergence free.

We have

$$\det\left(D\psi^1|D\psi^2\right) = \langle D\psi^1, D^\perp\psi^2\rangle,$$

where $D^\perp\psi^2 = (-\partial_y\psi^1, \partial_x\psi^1)$. Observe that $\operatorname{div}(D^\perp\psi^2) = 0$ (direct computation). By an integration by parts we then have

$$\int_{B^2} \det\left(D\psi^1|D\psi^2\right) = \int_{B^2} \langle D\psi^1, D^\perp\psi^2\rangle = \int_{\partial\mathbb{B}^2} \underbrace{\psi^1}_{=0} \nu\cdot D^\perp\psi^2 - \int_{B^2} \psi^1 \underbrace{\operatorname{div}(D^\perp\psi^2)}_{=0} = 0.$$

(One can also use Stokes' theorem to show this using differential forms).

That is,

$$\deg(\varphi) := \int_{B^n} \det(D\Phi),$$

is well defined, meaning it does not matter what the precise choice of $\Phi : \mathbb{B}^n \to \mathbb{R}^n$ is as long as $\Phi\Big|_{\partial\mathbb{B}^n} = \varphi$.

Now let $\varphi(x) = x$. Then, assuming the claim of the proposition is wrong, we find some $\Phi : \mathbb{B}^n \to \mathbb{S}^{n-1}$ that extends φ. But observe that this implies $\operatorname{rank} D\Phi \le n-1$ (otherwise by inverse function theorem $\Phi(B^n)$ is n-dimensional locally!). Thus $\det(D\Phi) = 0$. Thus

$$\deg(\varphi) = \int_{B^n} \det(D\Phi) = 0.$$

On the other hand we can choose $\Phi_2(x) := x$ and then we have

$$\deg(\varphi) = \int_{B^n} \det(Dx) = \int_{B^n} \det(I) = |B^n| \neq 0,$$

a contradiction, since the degree cannot be 0 and nonzero at the same time. □

Let us remark that degree theory from the proof above implies the famous

Theorem 14.2 (Brouwer Fixed Point)**.** *Every continuous map between the closed unit balls $\mathbb{B}^n$ has a fixed point. I.e. for any $f \in C^0(\mathbb{B}^n, \mathbb{B}^n)$ there exists $x_0 \in B^n$ such that $f(x_0) = x_0$.*

Let us remark that Theorem 14.2 can be extended relatively easily to convex compact sets.

Proof of Theorem 14.2. Assume this is not the case. Setting

$$\tilde{f}(x) := \begin{cases} f(x/|x|) & |x| \geq 1 \\ f(x) & |x| \leq 1 \end{cases}$$

we obtain a continuous map $\tilde{f} : \mathbb{R}^n \to \mathbb{B}^n$ without a fixed point.

By compactness of $\mathbb{B}^n$ we have

$$\lambda := \inf_{x \in \mathbb{B}^n} |f(x) - x| > 0.$$

If we now consider the usual mollification $\tilde{f}_\delta$ then we note that by convexity of $\mathbb{B}^n$ we have $\tilde{f}_\delta \in C^\infty(\mathbb{R}^n, \mathbb{B}^n)$, and by uniform convergence $\tilde{f}_\delta \to f$ we have for all small $\delta \ll 1$ and all $x \in \mathbb{B}^n$

$$|f_\delta(x) - x| \geq |f(x) - x| - \|f_\delta - f\|_{L^\infty(\mathbb{B}^n)} \geq \lambda - \underbrace{\|f_\delta - f\|_{L^\infty(\mathbb{B}^n)}}_{\ll 1} > 0.$$

That is, without loss of generality we can assume that $f \in C^\infty(\mathbb{B}^n, \mathbb{B}^n)$ and f has no fixed point.

For $t \in [0,1]$, $x \in \mathbb{B}^n$ we set

$$g_t(x) := \frac{x - tf(x)}{|x - tf(x)|},$$

For each $t \in [0,1]$ the map $g_t\Big|_{\mathbb{S}^{n-1}} \in C^\infty(\mathbb{S}^{n-1}, \mathbb{S}^{n-1})$. Indeed, we have more, even the dependency on t is C^∞-smooth:

$$(t,x) \ni [0,1] \times \mathbb{S}^{n-1} \mapsto g_t(x) \text{ is smooth.} \tag{14.1}$$

Indeed, the only obstacle to smoothness is the case when $|x - tf(x)| = 0$, but we have

$$\inf_{t \in [0,1]} \inf_{x \in \mathbb{S}^{n-1}} |x - tf(x)| > 0. \tag{14.2}$$

To see (14.2), observe that $\lambda := \inf_{\mathbb{B}^n} |f(x) - x| > 0$. So we have for any $t \in (1 - \frac{\lambda}{2}, 1]$

$$\inf_{x \in \mathbb{B}^n} |tf(x) - x| \geq \inf_{x \in \mathbb{B}^n} |f(x) - x| - |t-1| \sup_{x \in \mathbb{B}^n} |f(x)| \geq \lambda - |t-1| > \frac{\lambda}{2} > 0.$$

On the other hand for $(t,x) \in [0, 1 - \frac{\lambda}{4}] \times \mathbb{S}^{n-1}$ that $|x - tf(x)| \geq 1 - t \geq \frac{\lambda}{4}$. This implies (14.2), (14.1).

Observe moreover, that the above argument shows for $t = 1$ that the map $g_1 \in C^\infty(\mathbb{B}^n, \mathbb{S}^n)$ – since there is no fixed point.

If we now use the definition of degree from the proof of Proposition 14.1, we see that (observe $\det(Dg_1) = 0$ since g_1 maps into $\mathbb{S}^{n-1}$ by Inverse Function theorem)

$$\deg(g_1) = 0, \quad \text{and} \quad \deg(g_0) = \deg(\frac{x}{|x|}) = \int_{\mathbb{B}^n} \det(Dx) = 1. \tag{14.3}$$

So all we need to show is that $\underline{\deg(g_1) = \deg(g_0)}$ to get a contradiction.

For this denote let $G_0 \in C^\infty(\mathbb{B}^n, \mathbb{R}^n)$ be any smooth extension of g_0. Set now

$$H(x) := \begin{cases} g_{2|x|-1}(x/|x|) & |x| \in [\frac{1}{2}, 1] \\ G_0(2x) & |x| \le \frac{1}{2}. \end{cases}$$

Observe that $H \in C^0(\mathbb{B}^n, \mathbb{R}^n)$ – and by after a freezing argument and mollification we may assume that $H \in C^\infty(\mathbb{B}^n, \mathbb{R}^n)$ with the following properties

$$H(x) = \begin{cases} g_1(x/|x|) & |x| = 1 \\ g_0(x/|x|) & |x| = \frac{1}{2} \\ \in \mathbb{S}^{n-1} & |x| \in [\frac{1}{2}, 1]. \end{cases}$$

By the definition of degree we have

$$\deg(g_1) = \int_{\mathbb{B}^n} \det(DH)$$

and (a substituion argument is used here)

$$\deg(g_0) = \int_{\mathbb{B}^n} \det(D(H(\frac{1}{2}\cdot))) = \int_{\frac{1}{2}\mathbb{B}^n} \det(DH).$$

However, since $\operatorname{rank} DH(z) \le n-1$ for $\frac{1}{2} \le |z| \le 1$ we have

$$0 = \deg(g_1) = \int_{\mathbb{B}^n} \det(DH) = \int_{\frac{1}{2}\mathbb{B}^n} \det(DH) = \deg(g_0) = 1$$

This is a contradiction, so we can conclude. □

By approximation one can improve also Proposition 14.1 to continuous maps.

Proposition 14.3. *There is no continuous map $\Phi : \overline{\mathbb{B}^n} \to \mathbb{S}^{n-1}$ such that*

$$\Phi(x) = x \quad \forall x \in \mathbb{S}^{n-1}.$$

Proof. Assume there is. Then we can extend it to a continuous map $\tilde{\Phi} : \mathbb{R}^n \to \mathbb{S}^{n-1}$

$$\tilde{\Phi}(x) := \begin{cases} \frac{x}{|x|} & |x| \ge 1 \\ \Phi(x) & |x| < 1, \end{cases}$$

In view of Exercise 4.39 we can approximate $\tilde{\Phi}$ by a smooth function $\tilde{\Phi}_k \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ with $\tilde{\Phi}_k \equiv \frac{x}{|x|}$ for all $|x| \ge 2$ (since $\tilde{\Phi}$ is smooth in $|x| > \frac{3}{2}$ it is easy to check that the approximation in Exercise 4.39 is smooth.

This approximation is w.r.t. L^∞, so we have uniform convergence, so if we fix a large enough $k \gg 1$ we have $\|\tilde{\Phi}_k - \Phi\|_{L^\infty} < \frac{1}{2}$. Since $|\Phi(x)| \equiv 1$ we conclude $|\tilde{\Phi}_k| > \frac{1}{2}$ and thus

$$\tilde{\psi} := \frac{\tilde{\Phi}_k}{|\Phi_k|}$$

is a well-defined map in $C^\infty(\mathbb{R}^n, \mathbb{S}^{n-1})$ that satisfies $\tilde{\psi}(x) = \frac{x}{|x|}$ for all $|x| \ge 2$.

Considering $\psi(x) := \psi(2x)$ have found a counterexample to Proposition 14.1. $\square$

A pure reformulation of Proposition 14.3 is

Corollary 14.4. *Any map $f : \overline{\mathbb{B}^n(0,1)} \to \mathbb{S}^{n-1}$ which is the identity on $\mathbb{S}^{n-1} = \partial\mathbb{B}^n(0,1)$ is discontinuous.*

When working with Sobolev spaces, Proposition 14.3 readily implies

Corollary 14.5. *(1) Whenever $p > n$ there is no map $\Phi \in W^{1,p}(\mathbb{B}^n, \mathbb{S}^{n-1})$ such that – in the trace sense.*

$$\Phi(x) = x \quad \forall x \in \mathbb{S}^{n-1},$$

(2) Whenever $p < n$ there is a map $\Phi \in W^{1,p}(\mathbb{B}^n, \mathbb{S}^{n-1})$ such that – in the trace sense.

$$\Phi(x) = x \quad \forall x \in \mathbb{S}^{n-1}.$$

Proof. (1) Obvious from Proposition 14.3, since $W^{1,p}$-maps are continuous if $p > n$.
(2) Exercise 13.6.

$\square$

We can also treat the limit case. Keep in mind $W^{1,n}$-maps do not need to be continuous. $\log\log|x|$ is the typical example, Exercise 13.5.

Theorem 14.6. *There is no map $\Phi \in W^{1,n}(\mathbb{B}^n, \mathbb{S}^{n-1})$ such that .*

$$\Phi(x) = x \quad \forall x \in \mathbb{S}^{n-1},$$

(in the trace sense).

Proof. W.l.o.g. Φ is defined in $\mathbb{R}^n$ as follows

$$\Phi(x) := \begin{cases} \frac{x}{|x|} & |x| \geq 1 \\ \Phi(x) & |x| < 1, \end{cases}$$

since $\frac{x}{|x|}$ is smooth outside of zero and the traces coincide. We mollify Φ as in Exercise 13.20

$$\Phi_\delta(x) := \int_{\mathbb{R}^n} \eta(z)\Phi(x + \delta\theta(x)z)dz,$$

for some choice of $\theta \in C_c^\infty(B(0,2),[0,1])$ and $\theta \equiv 1$ in $B(0,3/2)$, and a typical bump function $\eta \in C_c^\infty(B(0,2))$, $\eta \equiv 1$ in $B(0,1)$ and $\int \eta = 1$.

Then for $\delta \ll 1$, Φ_δ is smooth, and $x/|x|$ outside of $B(0,2)$, so the only thing we need to ensure is that Φ maps close enough to the sphere.

Fix $x \in \mathbb{R}^n$. Observe that

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \leq |\Phi_\delta(x) - \Phi(y)| \quad \forall y \in \mathbb{R}^n. \tag{14.4}$$

If $\theta(x) = 0$ we can choose $y = x$ since then $\Phi_\delta(x) = \Phi(x)$, so in that case $\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) = 0$.

If $\theta(x) > 0$ set $\tilde{\delta} := \theta(x)\delta$. Integrating (14.4) with respect to y in $B(x, 2\tilde{\delta})$ we find

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \le \fint_{B(x,2\tilde{\delta})} |\Phi_\delta(x) - \Phi(y)| dy.$$

Now,

$$\Phi_\delta(x) = \int_{B(x,2\tilde{\delta})} \delta^{-n}\eta(z/\delta)\Phi(z)dz.$$

Thus (recall $\int \eta = 1$)

$$|\Phi_\delta(x) - \Phi(y)| \le C \fint_{B(x,2\tilde{\delta})} |\Phi(z) - \Phi(y)|dz,$$

that is

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \le C \fint_{B(x,2\tilde{\delta})} \fint_{B(x,2\tilde{\delta})} |\Phi(z) - \Phi(y)|dz\, dy.$$

Triangular inequality then shows

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \le C \fint_{B(x,2\tilde{\delta})} |\Phi(z) - (\Phi)_{B(x,2\tilde{\delta})}|dz.$$

Poincaré inequality and Hölder's inequality (observe the power of δ!)

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \le C\|\nabla\Phi\|_{L^n(B(x,2\tilde{\delta}))}.$$

By absolute continuity of the integral (since $\theta \in [0, 1]$ it does not play a role here) there exists a $\delta \ll 1$ such that

$$C\|\nabla\Phi\|_{L^n(B(x,2\tilde{\delta}))} < \frac{1}{2}.$$

That is for all $\delta \ll 1$,

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) < \frac{1}{2},$$

thus (up to scaling)

$$\frac{\Phi_\delta}{|\Phi_\delta|} : \mathbb{R}^n \to \mathbb{S}^{n-1}$$

is a smooth counterexample to Proposition 14.1. □

We can also make Proposition 14.3 more stable, relaxing the condition $\Phi(x) = x$ to $\Phi(x) \approx x$ on $\partial\mathbb{B}^n$.

Proposition 14.7. *There exists $\varepsilon > 0$ such that the following holds.*

There is no continuous map $\Phi : \overline{\mathbb{B}^n} \to \mathbb{S}^{n-1}$ such that

$$\sup_{x\in\mathbb{S}^{n-1}} |\Phi(x) - x| < \varepsilon. \tag{14.5}$$

Proof. As we shall see the constant ε is not very small, one should observe it simply depends on the "tubular neighborhood" of $\mathbb{S}^{n-1}$ where the projection exists.

As before, by contradiction and using mollification, we may assume that we have a counterexample $\Phi \in C^\infty(\mathbb{B}^n, \mathbb{S}^{n-1})$.

We define a *homotopy*

$$H(t,x) : [0,1] \times \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$$

by

$$H(t,x) := \frac{t\Phi(x) + (1-t)x}{|t\Phi(x) + (1-t)x|}$$

This makes sense since by Equation (14.5), $|t\Phi(x)+(1-t)x| \geq |x|-t|\Phi(x)-x| \geq 1-t\varepsilon > 0$ if ε is small enough. Clearly $H(t,x)$ is smooth in t and x. As in the proof of Theorem 14.2 we can now obtain that

$$\deg(H(1,\cdot)) = \deg((\Phi(\cdot)) = \int_{\mathbb{B}^n} \det(D\Phi) = 0$$

since rank $D\Phi \leq n-1$. Moreover

$$\deg(H(0,\cdot)) = \deg(x) = \int_{\mathbb{B}^n} 1 = |\mathbb{B}^n| > 0.$$

And also as in the proof of Theorem 14.2 we see that

$$t \mapsto \deg(H(t,\cdot))$$

must be constant. So we found the desired contradiction, and can conclude. □

14.2. **Sobolev spaces for maps between manifolds and the H=W problem.** We know that on reasonable (i.e. smoothly bounded) sets we can approximate any Sobolev functions by smooth functions, Theorem 13.15.

Let $1 \leq p < \infty$ and $\Omega \subset \mathbb{R}^n$ be a smoothly bounded set (or $\mathbb{R}^n$). It is very easy to define the Sobolev space $W^{k,p}(\Omega, \mathbb{R}^m)$.

$$W^{k,p}(\Omega, \mathbb{R}^m) := \{f = (f_1, \ldots, f_m) : \quad f \in W^{k,p}(\Omega)\}.$$

If we define

$$H^{k,p}(\Omega, \mathbb{R}^m\} := \overline{C^\infty(\overline{\Omega}, \mathbb{R}^m)}^{\|\cdot\|_{W^{k,p}(\mathbb{R}^n)}}$$

we get exactly the same space, namely

Exercise 14.8. *Let $1 \leq p < \infty$. Show that*

(1) *for any $f \in W^{k,p}(\mathbb{R}^n, \mathbb{R}^m)$ there exists $f_j \in C^\infty(\overline{\Omega}, \mathbb{R}^m)$ such that $\|f_j - f\|_{W^{k,p}(\mathbb{R}^n)} \xrightarrow{j\to\infty} 0$.*
(2) *Whenever $f \in C^\infty(\overline{\Omega}, \mathbb{R}^m)$ is a Cauchy sequence w.r.t. $W^{k,p}$-norm, then $f \in W^{k,p}(\Omega, \mathbb{R}^m)$.*

Now let $\mathcal{N}$ be a compact smooth d-manifold without boundary in $\mathbb{R}^n$ (observe this is equivalent to the above notion if instead $\mathcal{N}$ is an open set with smooth boundary!).

Let $\varphi_i : \Omega_i \subset \mathbb{R}^d \to \mathcal{N} \subset \mathbb{R}^n$ be any choice of parametrization (i.e. Ω_i are open and φ_i diffeomorphisms and $\bigcup_{i=1}^N \varphi_i(\Omega_i) = \mathcal{N}$. Then

$$W^{k,p}(\mathcal{N}, \mathbb{R}^m) := \left\{f : \mathcal{N} \to \mathbb{R}^m : \quad f \circ \varphi_i \in W^{k,p}(\Omega_i), \quad \forall i = 1, \ldots, N\right\}$$

We equip $W^{k,p}(\mathcal{N}, \mathbb{R}^m)$ (which is still a linear space) with the norm

$$\|f\|_{W^{k,p}(\mathcal{N},\mathbb{R}^m)} := \max_{i=1,\ldots,N} \|f \circ \varphi_i\|_{W^{k,p}(\Omega_i,\mathbb{R}^m)}$$

Exercise 14.9. *Show that in the above definition, the specific choice of φ_i does not matter. That is if $(\psi_j)_{j=1}^M$ is another choice of parametrization, then*

$$\begin{aligned}&\left\{f : \mathcal{N} \to \mathbb{R}^m : \quad f \circ \varphi_i \in W^{k,p}(\Omega_i), \quad \forall i = 1, \ldots, N\right\}\\ =&\left\{f : \mathcal{N} \to \mathbb{R}^m : \quad f \circ \psi_j \in W^{k,p}(\Omega_i), \quad \forall j = 1, \ldots, M\right\}.\end{aligned}$$

and the two norms are comparable.

$$\max_{i=1,\ldots,N} \|f \circ \varphi_i\|_{W^{k,p}(\Omega_i,\mathbb{R}^m)} \approx \max_{j=1,\ldots,M} \|f \circ \psi_j\|_{W^{k,p}(\Omega_j,\mathbb{R}^m)}$$

Hint: Exercise 13.17.

On the other hand, from Advanced Calculus we know what $f \in C^\infty(\mathcal{N}, \mathbb{R}^m)$ means: $f \circ \varphi_i \in C^\infty(\Omega_i, \mathbb{R}^m)$ for all i. So we can define again

$$H^{k,p}(\mathcal{N}, \mathbb{R}^m\} := \overline{C^\infty(\mathcal{N}, \mathbb{R}^m)}^{\|\cdot\|_{W^{k,p}(\mathbb{R}^n)}}$$

Again, the distinction between H and W is unneccesary, namely we have

Exercise 14.10. *Show that $H^{k,p}(\mathcal{N}, \mathbb{R}^m\} = W^{k,p}(\mathcal{N}, \mathbb{R}^m)$.*

Now let us restrict the target. If $\mathcal{M} \subset \mathbb{R}^m$ is a manifold, we define $W^{k,p}(\mathcal{N}, \mathcal{M})$ by restriction,

$$W^{k,p}(\mathcal{N}, \mathcal{M}) := \{f \in W^{k,p}(\mathcal{N}, \mathbb{R}^m) : \quad f(x) \text{ in } \mathcal{M} \text{ a.e. in } \mathcal{N}\}.$$

On the other hand we define again $H^{k,p}$ by approximation

$$H^{k,p}(\mathcal{N}, \mathcal{M}) := \overline{C^\infty(\mathcal{N}, \mathcal{M})}^{\|\cdot\|_{W^{k,p}(\mathcal{N},\mathbb{R}^m)}}$$

Exercise 14.11. *Let $\mathcal{N} \subset \mathbb{R}^N$ be a smooth, compact n-dimensional manifold without boundary. Show the respective versions of Theorem 13.46 and Theorem 13.49.*

Exercise 14.12. *Show for any $k = 0, 1, \ldots$*

$$H^{k,p}(\mathcal{N}, \mathcal{M}) \subset W^{k,p}(\mathcal{N}, \mathcal{M}).$$

The other inclusion is way more difficult – and we shall show: not always true! But let us first consider cases where it is true.

For this we need the following result from differential geometry:

Lemma 14.13. *Let $\mathcal{M}$ be a smooth, compact manifold without boundary in $\mathbb{R}^m$. Then there exists $\varepsilon > 0$ and setting the tubular neighborhood*

$$B_\varepsilon(\mathcal{M}) := \{x \in \mathbb{R}^m : \quad \operatorname{dist}(x, \mathcal{M}) < \varepsilon\}.$$

and a smooth map $\pi : B_\varepsilon(\mathcal{M}) \to \mathcal{M}$ such that $\pi(p) = p$ for all $p \in \mathcal{M}$.

Proof. The proof is not terribly difficult [Simon, 1996, Section 2.12.3]. The idea is that one can show that there exists a tubular neighborhood such that the *nearest point projection* $\pi : B_\varepsilon(\mathcal{M}) \to \mathcal{M}$ exists, i.e.

$$\pi(p) := q, \quad \text{where } q \in \mathcal{M} \text{ is such that} \quad |q - p| = \inf_{\tilde{q} \in \mathcal{M}} |\tilde{q} - p|.$$

Since $\mathcal{M}$ is compact, such a q exists for any $p \in \mathbb{R}^n$, but if ε is chosen small enough and $p \in B_\varepsilon(\mathcal{M})$ then q is *unique*, and then one can show that the dependency of q on p is smooth.

Let us remark that the choice of some $\varepsilon > 0$ is generally necessary, there is no continuous map $f : \mathbb{R}^n \to \mathbb{S}^{n-1}$ with $f(x) = x$ for all $x \in \mathbb{S}^{n-1}$, Corollary 14.4. □

We will restrict to the case where the domain is a manifold, since it is simpler.

Proposition 14.14. *Let $\Omega \subset \mathbb{R}^n$ be an smoothly bounded open set and $\mathcal{M} \subset \mathbb{R}^m$ be a compact manifold without boundary. Then*

$$H^{1,p}(\Omega, \mathcal{M}) = W^{1,p}(\Omega, \mathcal{M}).$$

whenever $p \geq n$.

Proof. For $p > n$ we could use the continuity of $W^{1,p}$-maps, but for $p = n$ this is not true anymore. Cf. Exercise 13.5.

Let $f \in W^{1,p}(\Omega, \mathcal{M})$. By an extension argument we can assume that $W^{1,p}(\Omega')$ where $\overline{\Omega} \subset \Omega'$ where Ω' is an open set. We simply "freeze" f close to the boundary $\partial\Omega$ (which is compact, so we have the projection $\pi_{\partial\Omega}$

$$f(x) := \begin{cases} f(x) & x \in \Omega \\ f(\pi_{\partial\Omega}(x)) & x \in \mathbb{R}^n \setminus \Omega \text{ but close to } \Omega. \end{cases}$$

Since

$$s(x) := \begin{cases} x & x \in \Omega \\ \pi_{\partial\Omega}(x) & x \in \mathbb{R}^n \setminus \Omega \text{ but close to } \Omega. \end{cases}$$

is Lipschitz we see that $f \circ s$ belongs to $W^{1,p}$, Remark 13.11.

Now we can mollify $f \circ s$, namely for $\delta \ll 1$ we set

$$g_\delta := (f \circ s) * \eta_\delta,$$

for the typical bump function $\eta \in C_c^\infty(B(0,2))$, $\eta \equiv 1$ in $B(0,1)$, $\int \eta = 1$.

Fix $x \in \Omega$. Then

$$\text{dist}\,(g_\delta(x), \mathcal{M}) \leq |g_\delta(x) - g(y)| \quad \forall y \text{ close enough to } \Omega. \tag{14.6}$$

Integrating (14.4) with respect to y in $B(x, 2\delta)$ we find

$$\text{dist}\,(g_\delta(x), \mathcal{M}) \leq \fint_{B(x,2\delta} |g_\delta(x) - g(y)|dy.$$

Now,

$$g_\delta(x) = \int_{B(x,2\delta)} \delta^{-n}\eta(z/\delta)g(z)dz.$$

Thus (recall $\int \eta = 1$)

$$|g_\delta(x) - g(y)| \leq C \fint_{B(x,2\delta)} |g(z) - g(y)|dz,$$

that is

$$\text{dist}\,(g_\delta(x), \mathcal{M}) \leq C\fint_{B(x,2\delta)}\fint_{B(x,2\delta)} |g(z) - g(y)|dz\,dy.$$

Triangular inequality then shows

$$\text{dist}\,(g_\delta(x), \mathcal{M}) \leq C\fint_{B(x,2\delta)} |g(z) - (g)_{B(x,2\delta)}|dz.$$

Poincaré inequality and Hölder's inequality (observe the power of δ!)

$$\text{dist}\,(g_\delta(x), \mathcal{M}) \leq C\delta^{1-n+n\frac{p}{p-1}n}\|\nabla g\|_{L^p(B(x,2\delta))}.$$

Since $p \geq n$ we have $1 - n + n\frac{p}{p-1}n \geq 0$ and thus

$$\text{dist}\,(g_\delta(x), \mathcal{M}) \leq C\|\nabla g\|_{L^p(B(x,2\delta))}.$$

By absolute continuity[38] for any $\varepsilon > 0$ there exists a $\delta \ll 1$ such that

$$C\|\nabla g\|_{L^n(B(x,2\delta))} < \varepsilon.$$

That is for all $\delta \ll 1$,

$$\text{dist}\,(g_\delta(x), \mathcal{M}) < \varepsilon,$$

thus

$$\pi_{\mathcal{M}} g_\delta \in C^\infty(\overline{\Omega})$$

is a well-defined smooth map. Since g_δ converges in $W^{1,p}$ to g, it is easy to show ($\pi_{\mathcal{M}}$ is a Lipschitz map!) that $\pi_{\mathcal{M}} g_\delta$ converges in $W^{1,n}$ to $\pi_{\mathcal{M}} g \equiv g$.

[38] And here the proof stops working if $p < n$, because if $1 - n + n\frac{p-1}{p} < 0$ for $p < n$, absolute continuity cannot compensate for this negative power of δ!

That is, any $W^{1,p}$-map, $p \geq n$, is smoothly approximable, which is what we needed to show. $\square$

Exercise 14.15. *Show Proposition 14.14 for $W^{1,p}(\mathcal{N},\mathcal{M})$ where $\mathcal{N} \subset \mathbb{R}^n$ is a manifold of dimension d and $p \geq d$.*

However, *It turns out that in general $H^{1,p}(\mathcal{N},\mathcal{M}) \neq W^{1,p}(\mathcal{N},\mathcal{M})$* if p is less than the dimension of $\mathcal{N}$, see Section 14.4.

First we need Fubini's theorem for Sobolev spaces.

14.3. **Fubini theorem for Sobolev spaces.** The following is the Sobolev version of the slicing for L^p-functions in Proposition 4.6.

Theorem 14.16. *Let $u, u_i \in W^{1,p}(\mathbb{R}^n)$ and let $1 \leq p < \infty$ such that*

$$\|u - u_i\|_{W^{1,p}} \xrightarrow{i\to\infty} 0.$$

Denote points in the cube by $(t,x') \in \mathbb{R}^n$, $t \in \mathbb{R}$, $x \in \mathbb{R}^{n-1}$. Then[39] *for $\mathcal{L}^1$-almost every $t \in \mathbb{R}$*

$$u(t,\cdot) \in W^{1,p}\left(\mathbb{R}^{n-1}\right).$$

Moreover there is a subsequence u_{i_j} such that[40]

$$u_{i_j}(t,\cdot) - u(t,\cdot) \xrightarrow{j\to\infty} 0 \quad \textit{in } W^{1,p}\left(\mathbb{R}^{n-1}\right)$$

for almost every $t \in \mathbb{R}$

Proof. There exists $u_k \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|u_k - u\|_{W^{1,p}(\mathbb{R}^n)} \xrightarrow{k\to\infty}$$

By the usual Fubini's theorem we have

$$0 \xleftarrow{k\to\infty} \int_{\mathbb{R}^n} |\nabla u_k - \nabla u|^p = \int_{\mathbb{R}} \left(\int_{\mathbb{R}^{n-1}} |\nabla u_k(t,x) - \nabla u(t,x)|^p dx\right) dt$$

So if we set

$$F_k(t) := \left(\int_{\mathbb{R}^{n-1}} |\nabla u_k(t,x') - \nabla u(t,x')|^p dx'\right)$$

we have in view of Theorem 3.51 that for some subsequence F_{k_j}

$$F_{k_j}(t) \xrightarrow{j\to\infty} 0 \quad \mathcal{L}^1\text{-a.e. } t.$$

[39]Here a bit care is needed: u as a measurable map is only defined almost everywhere. What we say here is that each representative u can be restricted two a set $\Sigma \subset \mathbb{R}$ on which this representative is a representative of a Sobolev map. Two different representatives will lead to two different sets $\Sigma \subset \mathbb{R}$.

[40]Spoiler: this is the main point. Observe that Sobolev embedding depends on the dimension and $1 - \frac{n}{p} < 1 - \frac{n-1}{p}$. I.e. if $1 - \frac{n-1}{p} > 0$ then we have control of the continuity on $\mathbb{R}^{n-1}$, whereas we might not have any continuity in $\mathbb{R}^n$!

If we denote

$$\nabla' f(t,x') := (\partial_2 f, \ldots, \partial_n f)^t(t,x')$$

we find, observing that $|\nabla' f| \leq |\nabla f|$, that for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$

$$\int_{\mathbb{R}^{n-1}} |\nabla' u_k(t,x') - \nabla' u(t,x')|^p dx' \xrightarrow{k\to\infty} 0.$$

Similarly we can argue for the L^p-norm, so (after passing to a more refined subsequence labeled the same way) we have

$$\|u_{k_j}(t,\cdot) - u(t,\cdot)\|_{W^{1,p}(\mathbb{R}^{n-1})} \xrightarrow{j\to\infty} 0.$$

But since $u_{k_j}(t,\cdot) \in C_c^\infty(\mathbb{R}^{n-1})$ for all $t \in \mathbb{R}$ we conclude that this means its limit $u(t,\cdot) \in W^{1,p}(\mathbb{R}^{n-1})$.

We can argue the same way for $(u_i)_{i\in\mathbb{N}}$. Each u_i can be approximated by some smooth $(u_{i;k})_k$ and by triangular inequality the convergence holds. □

Exercise 14.17. *Show the following: Let $u, u_i \in W^{1,p}(B(0,1))$ and let $1 \leq p < \infty$ such that*

$$\|u - u_i\|_{W^{1,p}} \xrightarrow{i\to\infty} 0.$$

Every point in $x \in B(0,1)$ besides 0 can be uniquely written as $x = r\theta$ where $r \in (0,1)$ and $\theta \in \mathbb{S}^{n-1}$.

For $\mathcal{L}^1$-a.e. $r \in (0,1)$ we have

$$u(r\cdot) \in W^{1,p}\left(\mathbb{S}^{n-1}\right).$$

Moreover there is a subsequence u_{i_j} such that

$$u_{i_j}(r\cdot) - u(r\cdot) \xrightarrow{j\to\infty} \quad \text{in } W^{1,p}\left(\mathbb{S}^{n-1}\right)$$

for $\mathcal{L}^1$-almost every $r \in (0,1)$

Exercise 14.18. *Show Theorem 14.16 for $W^{m,p}$.*

14.4. **Nondensity of smooth maps in Sobolev spaces between manifolds.** As an application of Fubini's theorem for Sobolev spaces, Section 14.3, we answer (negatively) the $H = W$ question for Sobolev maps between manifold from Section 14.2

Theorem 14.19. *Let $n-1 < p < n$. There exists a map $u \in W^{1,p}(B^n(0,1), \mathbb{S}^{n-1})$ that cannot be smoothly approximated by maps $u_k \in C^\infty(\overline{B^n(0,1)}, \mathbb{S}^{n-1})$*

Before we come to the proof of Theorem 14.19, observe this is a special feature of Sobolev classes, which we have not seen e.g. for continuous classes. Namely we have

Exercise 14.20. *Let $1 < p < \infty$ and $\mathcal{M}$ any smooth compact manifold without boundary. Then any $u \in C^0 \cap W^{1,p}(B^n(0,1), \mathcal{M})$ can be smoothly approximated by maps into the by maps $u_k \in C^\infty(\overline{B^n(0,1)}, \mathcal{M})$*

Hint: mollify to obtain $u_\delta \in C^\infty(B^n(0,1), \mathbb{R}^n)$ and then control $\operatorname{dist}(u_\delta, \mathcal{M})$ by uniform convergence.

Proof of Theorem 14.19. Take $u(x) := \frac{x}{|x|}$. By Exercise 13.6 $u \in W^{1,p}(B(0,1), \mathbb{S}^{n-1})$.

We claim that u cannot be smoothly approximated.

Assume by contradiction, there exists a sequence $u_k \in C^\infty(B(0,1), \mathbb{S}^{n-1})$ such that $\|u_k - u\|_{W^{1,p}(B(0,1))} \xrightarrow{k\to\infty} 0$.

By Fubini's theorem, Exercise 14.17, up to passing to a subsequence, we can find some $r \in (0,1)$ such that $f_k(x) := u_k(rx)$ converges to $f(x) := u(rx) = \frac{x}{|x|}$ in $W^{1,p}(\partial\mathbb{S}^{n-1})$.

By Sobolev-Morrey embedding, Exercise 14.11, since $p > n-1$ (which is the dimension of $\mathbb{S}^{n-1}$), we have that f_k converges uniformly to x on $\mathbb{S}^{n-1}$,

$$\|f_k - x\|_{L^\infty(\mathbb{S}^{n-1})} \xrightarrow{k\to\infty} 0.$$

By Proposition 14.7 this cannot be true, contradiction. □

The general theory of these topological obstructions is due to Bethuel [Bethuel, 1991] with corrections by Hang and Lin [Hang and Lin, 2001], and is related with many open questions [Bethuel, 2020].

15. BV

When defining $W^{1,p}$ we observed that for $1 < p < \infty$ weak closure (i.e. closure under weak topology) and strong closure (closure under strong topology) of C^∞ under the $W^{1,p}$-norm give the same space.

A particular effect of this is:

Exercise 15.1. *Let $\Omega \subset \mathbb{R}^n$ be smooth and bounded. Let $p \in (1,\infty]$ and assume that $f_n \xrightarrow{n\to\infty} f$ almost everywhere in Ω*

Show that

$$\|f\|_{W^{1,p}(\Omega)} \leq \liminf_{n\to\infty} \|f_n\|_{W^{1,p}(\Omega)}.$$

whenever the right-hand side is finite.

Hint: For $p \in (1,\infty)$ use reflexivity and lower semicontinuity of the norm, as well as Rellich's theorem. If $p = \infty$ use Arzela-Ascoli and the identification of Sobolev and Lipschitz spaces.

This is not the case for $p = 1$.

Example 15.2. Let $\Omega = (-1, 1)$ and set

$$u_n(x) = \begin{cases} 0 & -1 < x < 0 \\ nx & 0 < x < \frac{1}{n} \\ 1 & x > \frac{1}{n}. \end{cases}$$

Each u_n is Lipschitz continuous so in particular $u_n \in W^{1,1}((-1,1))$, and

$$\|u_n\|_{L^1(-1,1)} \leq 2,$$

and

$$\|u_n'\|_{L^1(-1,1)} = \int_0^{\frac{1}{n}} n dx = 1.$$

Set

$$u(x) = \begin{cases} 0 & -1 < x < 0 \\ 1 & x > 0. \end{cases}$$

Then u is discontinuous and thus (we are in one dimension!) $u \notin W^{1,1}$. However we see

$$\|u_n - u\|_{L^1((-1,1))} \xrightarrow{n\to\infty} 0,$$

e.g. by the dominated convergence theorem.

So

$$\|u\|_{W^{1,1}} \not\leq \liminf_{n\to\infty} \|u_n\|_{W^{1,1}(\mathbb{R}^n)}.$$

The strong closure of C^∞ under the $W^{1,1}$-norm gives $W^{1,1}$ (functions in L^1 whose distributional derivative belongs to L^1).

But the weak closure leads to the space of bounded variations, $f \in L^1$ such that Df is a measure, cf. Exercise 12.40. Since we want to avoid the use of signed measures, we will first define BV by duality.

First let us consider the $W^{1,p}$-version. Recall that for $\phi \in C_c^1(\Omega, \mathbb{R}^n)$ the *divergence* $\operatorname{div} \phi \in C_c^0(\Omega)$ is given by

$$\operatorname{div} \phi = \sum_{i=1}^n \partial_i \phi^i.$$

Theorem 15.3. *Let $\Omega \subset \mathbb{R}^n$ be open and smoothly bounded (for simplicity), $p \in (1, \infty)$, and $f \in L^p(\Omega)$. Then $f \in W^{1,p}(\Omega)$ if and only if*

$$\Lambda := \sup \left\{ \int_\Omega f \operatorname{div} \phi d\mathcal{L}^n : \quad \phi \in C_c^1(\Omega, \mathbb{R}^n), \|\phi\|_{L^{p'}(\Omega)} \leq 1 \right\} < \infty.$$

In that case there exists a constant such that $C^{-1}\|Df\|_{L^p} \leq \Lambda \leq C\|Df\|_{L^p}$.

Proof. $\Rightarrow$ Assume $f \in W^{1,p}(\Omega)$ then

$$\int f \operatorname{div} \phi = \int Df \cdot \phi \le \|Df\|_{L^p} \|\phi\|_{L^{p'}}$$

$\Leftarrow$ Let $\phi = (0, \dots, 0, \varphi, 0, \dots, 0)$ (where φ is at the j-th position then we get that

$$\varphi \mapsto \int_\Omega f \partial_j \varphi \le \Lambda \|\varphi\|_{L^{p'}(\mathbb{R}^n)}.$$

By the Riesz representation theorem (here we need that $p' < \infty$, so $p > 1$) we find $g \in L^p(\mathbb{R}^n)$, $\|g\|_{L^p} \le \Lambda$ such that

$$\int_\Omega f \partial_j \varphi = \int_\Omega g \varphi,$$

that is $f \in W^{1,p}$. □

Motivated by this we define BV as the $p = 1$ case of the above theorem.

Definition 15.4. Let $\Omega \subset \mathbb{R}^n$ be open.

(1) A function $f \in L^1(\Omega)$ is said to have *bounded variation* in Ω if

$$\sup \left\{ \int_\Omega f \operatorname{div} \phi d\mathcal{L}^n : \quad \phi \in C_c^1(\Omega, \mathbb{R}^n), |\phi(x)| \le 1 \text{ in } \Omega \right\} < \infty.$$

(2) the collection of functions with bounded variations is denoted by $BV(\Omega)$. *Here we do not identify two functions which agree $\mathcal{L}^n$-a.e.*.
(3) An $\mathcal{L}^n$-measurable subset $E \subset \mathbb{R}^n$ has *finite perimeter* in Ω if $\chi_E \in BV(\Omega)$.
(4) A function $f \in L^1_{loc}(\Omega)$ is said to have *locally bounded variation* in Ω if for each $\Omega' \subset\subset \Omega$,

$$\sup \left\{ \int_{\Omega'} f \operatorname{div} \phi d\mathcal{L}^n : \quad \phi \in C_c^1(\Omega', \mathbb{R}^n), |\phi(x)| \le 1 \text{ in } \Omega' \right\} < \infty.$$

(5) the collection of functions with locally bounded variations is denoted by $BV_{loc}(\Omega)$. *Here we do not identify two functions which agree $\mathcal{L}^n$-a.e.*.
(6) An $\mathcal{L}^n$-measurable subset $E \subset \mathbb{R}^n$ has *locally finite perimeter* in Ω if $\chi_E \in BV_{loc}(\Omega)$ (such a set is sometimes called a *Caccioppoli set*)

As mentioned, BV means that Df is a Radon measure. The precise meaning of that is the following

Theorem 15.5. *Assume that $f \in BV_{loc}(\Omega)$.*

Then there exists a (nonnegative) Radon measure μ on Ω and a μ-measurable function $\sigma : \Omega \to \mathbb{R}^n$ such that

(1) $|\sigma(x)| = 1$ for μ-a.e. x, and

(2) for all $\phi \in C_c^1(\Omega;\mathbb{R}^n)$ we have
$$\int_\Omega f \operatorname{div}\phi = -\int_\Omega \phi\cdot\sigma\, d\mu.$$

Proof. We skip the proof, it is relatively straight-forward consequence of the Riesz Representation theorem Theorem 5.44, similar to $W^{1,p}$-case above. □

So $\sigma\llcorner\mu$ takes the role of Df. We will write
$$\|Df\| \quad \text{for the measure } \mu, \quad \text{the \textit{variation measure}}$$
and
$$[Df] := \|Df\|\llcorner\sigma$$
so that we have in the above theorem
$$\int_\Omega f \operatorname{div}\phi = -\int_\Omega \phi\cdot\sigma d\|Df\| \equiv -\int_\Omega \phi\cdot d[Df] \quad \forall \phi\in C_c^1(\Omega;\mathbb{R}^n).$$
In a similar spirit, if $f=\chi_E$ and E is a set of locally finite perimeter in U we write
$$\nu_E := -\sigma$$
and
$$\|\partial E\| := \mu \quad \text{the \textit{perimeter measure}}$$
so that
$$\int_E \operatorname{div}\phi dx = \int_U \phi\cdot\nu_E\; d\|\partial E\| \quad \forall\phi\in C_c^1(\Omega;\mathbb{R}^n).$$
We furthermore can apply Lebesgue decomposition theorem, Theorem 5.14, and split
$$\|Df\| = \|Df\|_{ac} + \|Df\|_s,$$
where $\|Df\|_{ac} \ll \mathcal{L}^n$ and $\|Df\|_s \perp \mathcal{L}^n$. We then have for some $G\in L^1_{loc}(U,\mathbb{R}^n)$
$$\sigma\|Df\|_{ac}(\Omega) = \int_\Omega G(x)\, d\mathcal{L}^n(x).$$
We denote $Df(x) := G(x)$ the density of the absolute continuous part of $[Df]$. Then we have
$$[Df] = \mathcal{L}^n\llcorner Df + [Df]_s.$$

If $f\in BV_{loc}(U)\cap L^1(U)$ then $f\in BV_{loc}(U)$ if and only if $\|Df\|(U)<\infty$, and in this case we define
$$\|f\|_{BV(U)} := \|f\|_{L^1(U)} + \|Df\|(U).$$
Observe that from the Riesz representation theorem we have
$$\|Df\|(\Omega) := \sup\left\{\int_\Omega f \operatorname{div}\phi d\mathcal{L}^n : \quad \phi\in C_c^1(\Omega,\mathbb{R}^n), |\phi(x)|\leq 1 \text{ in } \Omega\right\},$$
and
$$\|\partial E\|(\Omega) := \sup\left\{\int_\Omega \chi_E \operatorname{div}\phi d\mathcal{L}^n : \quad \phi\in C_c^1(\Omega,\mathbb{R}^n), |\phi(x)|\leq 1 \text{ in } \Omega\right\}.$$

Lemma 15.6. • *$f \in W^{1,1}_{loc}(U)$ then $f \in BV_{loc}(U)$*
- *$f \in W^{1,1}(U)$ then $f \in BV(U)$*

Proof. This follows from the integration by parts formula for $\phi \in C^1_c(\Omega, \mathbb{R}^n)$

$$\int_\Omega f \operatorname{div} \phi = \int_\Omega \nabla f \cdot \phi \le \|\phi\|_{L^\infty} \|\nabla f\|_{L^1}.$$

□

Definition 15.7. For a set $E \subset \mathbb{R}^n$ of finite perimeter we set

$$\operatorname{Per}(E) := \|\partial E\|(\mathbb{R}^n)$$

the *perimeter* of E.

Example 15.8. Assume E is a smooth open subset of $\mathbb{R}^n$ and $\mathcal{H}^{n-1}(\partial E) < \infty$ then we have from the integration by parts formula

$$\int_E \operatorname{div}(\phi) = \int_{\partial E} \nu \cdot \phi \le \mathcal{H}^{n-1}(\partial E) \, \|\phi\|_{L^\infty(\partial E)} \le \mathcal{H}^{n-1}(\partial E) \|\phi\|_{L^\infty(\mathbb{R}^n)} \quad \forall \phi \in C^1_c(\mathbb{R}^n).$$

That is $\chi_E \in BV_{loc}(\mathbb{R}^n)$. But observe that $\chi_E \not\in W^{1,1}_{loc}(\mathbb{R}^n)$. Indeed if $\chi_E \in W^{1,1}_{loc}(\mathbb{R}^n)$, then using Fubini's theorem Section 14.3 iteratively on smaller and smaller dimensions, there would be some straight line L intersecting E such that $\chi_E \in W^{1,1}(L)$ – but then χ_E would have a continuous representative on that straigt line Proposition 13.50, which is impossible since it jumps from 1 to 0.

Example 15.8 shows also that for all sufficiently smooth sets E, we have

$$\operatorname{Per}(E) = \mathcal{H}^{n-1}(\partial E).$$

For general sets of finite perimeter this will not be true (otherwise the notion of Perimeter would be quite useless, wouldn't it?) but it is true for the so-caelled *reduced boundary* $\partial^* E$, see Theorem 15.19.

15.1. **Some properties: Lower semicontinuity, Approximation, Compactness, traces.**

Theorem 15.9 (Lower semicontinuity)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set. Suppose $(f_k)_{k\in\mathbb{N}} \subset BV(\Omega)$ and $f_k \to f \in L^1_{loc}(\Omega)$.*

$$\|Df\|(\Omega) \le \liminf_{k\to\infty} \|Df_k\|(\Omega),$$

In the sense that if $\liminf_{k\to\infty} \|Df_k\|(\Omega) < \infty$ then $f \in BV_{loc}(\Omega)$ and we have the above inequality.

Proof. Fix $\phi \in C_c^1(\Omega, \mathbb{R}^n)$. Then, by L^1-convergence

$$\begin{aligned}\int_\Omega f \operatorname{div}\phi &= \lim_{k\to\infty}\int_\Omega f_k \operatorname{div}\phi\\ &= \lim_{k\to\infty}\int_\Omega \phi\cdot\sigma_k d\|Df_k\|\\ &\leq \|\phi\|_{L^\infty} \liminf_{k\to\infty}\|Df_k\|(\Omega)\end{aligned}$$

□

Theorem 15.10. *Let $\Omega \subset \mathbb{R}^n$ be an open set. Assume $f \in BV(\Omega)$. Then there exist functions $(f_k)_{k\in\mathbb{N}} \subset BV(\Omega)\cap C^\infty(\Omega)$ (observe the openness of Ω we do not get continuity up to $\partial\Omega$) such that*

1. *$f_k \xrightarrow{k\to\infty} f$ in $L^1(\Omega)$*
2. *$\|Df_k\|(\Omega) \xrightarrow{k\to\infty} \|Df\|(\Omega)$.*
3. *If we denote $\mu_k(B) := \int_{B\cap\Omega} Df_k dx$ for each Borel set $B \subset \mathbb{R}^n$, and set $\mu(B) := \int_{B\cap U} d[Df]$ then μ_k weakly converges to μ in the sense of (vector-valued) Radon measures in $\mathbb{R}^n$.*

Proof. For the full proof we refer to [Evans and Gariepy, 2015, Theorem 5.3 and Theorem 5.4], however let us sketch the main idea here.

Let $\eta \in C_c^\infty(B(0,1))$, $\eta(x) = \eta(-x)$, $\int \eta = 1$ be the typical mollification bump function. It is very reasonable to hope that $f_\varepsilon := f * \varepsilon^{-n}\eta(\cdot/\varepsilon)$ is the right approximation for f – but there is the smearing out of the convolution. Nevertheless, from lower semicontinuity, Theorem 15.9, we have (where we consider f extended by zero outside of Ω)

$$\|Df\|(\Omega) \leq \liminf_{\varepsilon\to 0}\|Df_\varepsilon\|(\Omega)$$

So all we have to show is the opposite direction – which is messy, and here we are only going to show

$$\limsup_{\varepsilon\to 0}\|Df_\varepsilon\|(\Omega') \leq \|Df\|(\Omega),$$

for any open $\Omega' \subset\subset \Omega$ (i.e. $\Omega' \subset \overline{\Omega'} \subset \Omega$ and $\overline{\Omega'}$ is compact).

To see this take ε small enough so that $B_\varepsilon(\Omega') \subset \Omega$. Then for any $\Phi \in C_c^\infty(\Omega', \mathbb{R}^n)$, by Fubini's theorem and integration by parts (and since $\eta(x) = \eta(-x)$),

$$\int_{\Omega'} f_\varepsilon \operatorname{div}\Phi = \int_\Omega f \operatorname{div}(\Phi * \eta) \leq \|\Phi * \eta\|_{L^\infty}\|Df\|(\Omega) \leq \|\Phi\|_{L^\infty}\|Df\|(\Omega)$$

Taking the supremum over such Φ with $\|\Phi\|_{L^\infty} \leq 1$, we readily find

$$\|Df_\varepsilon\|(\Omega') \leq \|Df\|(\Omega) \quad \text{for all small } \varepsilon > 0,$$

that is,

$$\limsup_{\varepsilon\to 0}\|Df_\varepsilon\|(\Omega') \leq \|Df\|(\Omega).$$

The full proof is then a careful covering argument, but follows the spirit from the argument above. $\square$

Exercise 15.11. *Show that if $\Omega = \mathbb{R}^n$ and $f \in BV(\Omega)$ we can choose the approximation f_k in Theorem 15.10 (1) and (2) to belong to $f_k \in C_c^\infty(\mathbb{R}^n)$*

Remark 15.12. Observe

- there is no assumption on regularity on $\partial\Omega$.
- We do not claim that $\|Df_k - Df\|(U) \xrightarrow{k\to\infty} 0$ (because this would indicate that $Df \in L^1_{loc}$, and take $f = \chi_E$ as counterexample where E is a nice set)

We get a version of Rellich's theorem, Theorem 13.35.

Theorem 15.13 (Compactness / Rellich's theorem)**.** *Let $U \subset \mathbb{R}^n$ be open and bounded with Lipschitz boundary ∂U. Assume $(f_k)_{k\in\mathbb{N}}$ is a squeunce in $BV(U)$ satisfying*

$$\sup_k \|f_k\|_{BV(U)} < \infty.$$

Then there exists a subsequence $(f_{k_j})_{j\in\mathbb{N}}$ and a function $f \in BV(U)$ such that

$$f_{k_j} \xrightarrow{j\to\infty} f \quad in\ L^1(U).$$

Proof. Follows from Rellich's theorem, just approximate f_k by $g_k \in C^\infty$ (and thus $g_k \in W^{1,1}_{loc}$), such that

$$\|f_k - g_k\|_{L^1} < \frac{1}{k}$$

and

$$\|Dg_k\|_{L^1} = \|Dg_k\|(U) \leq \|Df_k\|(U) + \frac{1}{k}.$$

We conclude that (g_k) is bounded in $W^{1,1}$, by Rellich's theorem, Theorem 13.35, (there is a subsequence) g_{k_j} which converges strongly in $L^1(\mathcal{U})$, and so f_k converges in $L^1(U)$. $\square$

If $f \in W^{1,p}(\Omega)$ and Ω is bounded with smooth boundary, then $f\Big|_{\partial\Omega} \in W^{1-\frac{1}{p},p}$ (this works if $p \geq 1$, with $W^{0,1} = L^1$). Essentially this still works for BV-functions, in the following sense

Theorem 15.14 (Trace)**.** *Let U be open and bounded with ∂U Lipschitz continuous. There exists a bounded linear mapping*

$$T : BV(U) \to L^1(\partial U; \mathcal{H}^{n-1}),$$

such that

$$\int_U f \operatorname{div} \phi dx = -\int_U \phi \cdot d[Df] + \int_{\partial U} \phi \cdot \nu\, Tf d\mathcal{H}^{n-1}$$

for all $f \in BV(U)$ and $\phi \in C^1(\overline{U}, \mathbb{R}^n)$.

Tf is called the trace of f on ∂U – and if $f \in W^{1,1}(U) \subset BV(U)$ then it coincides with the trace of Theorem 13.31.

For a proof see [Evans and Gariepy, 2015, Theorem 5.6]

Theorem 15.15 (Sobolev embedding)**.** *Let $u \in BV(\mathbb{R}^n)$. Then*

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq C\,\|Du\|(\mathbb{R}^n)$$

where C is a constant only depending on the dimension n.

Proof. By Exercise 15.11 we can approximate u by $u_k \in C_c^\infty(\mathbb{R}^n)$ such that $\|u_k - u\|_{L^1(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$, $f_k \xrightarrow{k\to\infty} f$ $\mathcal{L}^n$-a.e., and $\|Df_k\|(\mathbb{R}^n) \xrightarrow{k\to\infty} \|Df\|(\mathbb{R}^n)$.

By Theorem 13.41, for each $k \in \mathbb{N}$

$$\|u_k\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq C\,\|Du_k\|_{L^1(\mathbb{R}^n)} = C\,\|Du_k\|(\mathbb{R}^n).$$

By Fatou's Lemma, Corollary 3.9,

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq \liminf_{k\to\infty} \|u_k\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq C\,\liminf_{k\to\infty} \|Du_k\|(\mathbb{R}^n) = C\,\|Du\|(\mathbb{R}^n).$$

□

15.2. **Isoperimetric inequality.** The isoperimetric inequality relates the *area* of a set E with the *length* of its boundary ∂E. In its simplest form (an easy consequence of Sobolev-inequality for BV, Theorem 15.15) it looks as follows

Theorem 15.16 (Isoperimetric inequality)**.** *Let $E \subset \mathbb{R}^n$ be a set of finite perimeter. Then*

$$\mathcal{L}^n(E)^{\frac{n-1}{n}} \leq C\mathrm{Per}(E)$$

Proof. In this form this is the Sobolev theorem Theorem 15.15, for χ_E, then we have

$$\|\chi_E\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq C\|D\chi_E\|(\mathbb{R}^n) \equiv \mathrm{Per(E)}.$$

Now we observe that

$$\|\chi_E\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} = \mathcal{L}^n(E)^{\frac{n-1}{n}}$$

□

It is very interesting to find out what the optimal constant C is, and what shape E needs to have for this optimal constant to be attained (ball!). This is caleld Dido's problem cf. [Bandle, 2017].

15.3. **Reduced boundaries.** Without going into too much detail let us discuss some cool features about sets with finite perimeter. For details and proofs see [Evans and Gariepy, 2015, Chapter 5].

For sets E of finite perimeter we have that ∂E can be quite a wild set. The measure $\|\partial E\|$ leads to another notion of boundary: the points where $\|\partial E\|$ has suitable density.

Definition 15.17. Let $E \subset \mathbb{R}^n$ be a set of locally finite perimeter. The *reduced boundary* $\partial^* E$ of E is defined as the collection of all $x \in \mathbb{R}^n$ such that

(1) $\|\partial E\|(B(x,r)) > 0$ for all $r > 0$, and
(2) $\lim_{r\to 0} \fint_{B(x,r)} \nu_E d\|\partial E\| = \nu_E(x)$, and
(3) $|\nu_E(x)| = 1$.

Lemma 15.18. *$\partial^* E \subset \partial E$.*

Proof. Indeed, this follows from the first condition. If $x \notin \partial E$ then there exists a radius $r > 0$ such that the ball $B(x,r) \subset E$ or $B(x,r) \subset \mathbb{R}^n \setminus E$. Take any $\phi \in C_c^\infty(B(x,r), \mathbb{R}^n)$ then we have

$$\int_E \operatorname{div} \phi = \int_{E\cap B(x,r)} \operatorname{div} \phi.$$

Either, if $B(x,r) \subset E$

$$\int_E \operatorname{div} \phi = \int_{B(x,r)} \operatorname{div} \phi = 0,$$

or if $B(x,r) \in \mathbb{R}^n \setminus E$ we have

$$\int_{E\cap B(x,r)} \operatorname{div} \phi = \int_\emptyset \operatorname{div} \phi = 0.$$

So in either case $\int_E \operatorname{div} \phi = 0$, so $\|\partial E\|(B(x,r)) = 0$. □

Theorem 15.19 (Structure theorem for sets of finite perimeter)**.** *Let $E \subset \mathbb{R}^n$ be a set of locally finite perimeter.*

- *$\mathcal{H}^{n-1}(B) \le C\|\partial E\|(B)$ for all $B \subset \partial^* E$ (where C is a constant depending only on n)*
- *The reduced boundary $\partial^* E$ can be written as*

$$\partial^* E = \bigcup_{k=1}^\infty K_k \cup N,$$

where

$$\|\partial E\|(N) = 0$$

and each K_k is a compact subset of a C^1-hypersurface S_k.
- *$\nu_E\Big|_{S_k}$ is normal to S_k for each k*
- *$\|\partial E\| = \mathcal{H}^{n-1} \llcorner \partial^* E$. In particular*

$$\operatorname{Per}(E) = \mathcal{H}^{n-1}(\partial^* E).$$

Definition 15.20 (measure-theoretic boundary)**.** Let $E \subset \mathbb{R}^n$. The *measure-theoretic boundary* $\partial_* E$ of E is given by all $x \in \mathbb{R}^n$ which satisfy

$$\limsup_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap E)}{r^n} > 0$$

and

$$\limsup_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\setminus E)}{r^n} > 0$$

We have $\partial^* E \subset \partial_* E$ and $\mathcal{H}^{n-1}(\partial_* E \setminus \partial^* E) = 0$.

Theorem 15.21 (Gauss-green)**.** *Let $E \subset \mathbb{R}^n$ have locally finite perimeter.*

(1) Then $\mathcal{H}^{n-1}(\partial_* E \cap K) < \infty$ for each compact set $K \subset \mathbb{R}^n$.
(2) For $\mathcal{H}^{n-1}$-a.e. $x \in \partial_* E$, there is a unique measure theoretic unit outer normal $\nu_E(x)$ (this of course needs to be defined) such that

$$\int_E \operatorname{div}\phi = \int_{\partial_* E} \phi\cdot \nu_E d\mathcal{H}^{n-1} \quad \forall \phi \in C^1_c(\mathbb{R}^n,\mathbb{R}^n).$$

One can show that if $E \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable and $\mathcal{H}^{n-1}(\partial_* E\cap K) < \infty$ for all compact sets $K \subset \mathbb{R}^n$, then E has locally finite paramter (in particular Theorem 15.21 holds for open sets with Lipschitz boundary)

Part 3. Analysis III: Cool Tools from Functional Analysis

16. Short crash course on main examples: L^p and $W^{1,p}$

We are going to use those a lot, so let us repeat their basic definitions.

Let $\Omega \subset \mathbb{R}^n$ be an open set (this will be a standing assumption).

For $p \in [1,\infty)$ we say $f \in L^p(\Omega)$ if

$$\|f\|_{L^p(\Omega)} := \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} < \infty.$$

This looks very simple, but there are a few things to observe here:

- This is the Lebesgue integral, which essentially means it is defined for a larger class of functions than if it was the Riemann integral (there are no other practical differences). Namely $f : \Omega \to \mathbb{R}$ needs to be *measurable* then $(\int_\Omega |f|^p)^{\frac{1}{p}}$ is a defined quantity (which could be infinite). Again, this has no real practical consequences, thinking of the Riemann integral is not too far of, namely:

- For every $f \in L^p(\Omega)$ there exists a sequence $f_k \in C^0(\overline{\Omega}) \cap L^p(\Omega)$ that converges to f, i.e.
$$\|f_k - f\|_{L^p(\Omega)} \xrightarrow{k\to\infty} 0,$$
and thus in particular
$$\|f\|_{L^p(\Omega)} = \lim_{k\to\infty} \left(\int_\Omega |f_k|^p\right)^{\frac{1}{p}}.$$
For continuous f_k Riemann and Lebesgue integral are the same – so thinking of $\int_\Omega$ as the usual Riemann integral is just fine (essentially).
- We call $\|\cdot\|_{L^p(\Omega)}$ a norm, technically it is not because $\|f\|_{L^p(\Omega)} = 0$ does not imply $f(x) = 0$ for all $x \in \Omega$ (take e.g. $f = 0$ in $\Omega \setminus \{x_0\}$ and $f(x_0) = 1$).

  The solution to this issue is that we consider two functions f and g to be the same (in the sense of integration theory) if
$$f(x) = g(x) \quad \text{for almost every } x \in \Omega,$$
where "almost all" (short: "a.e.") is a precise notion: it means that the set N of all x where $f(x) \neq g(x)$ is a set of Lebesgue measure zero, which means that for any $\varepsilon > 0$ there exists a countable ball cover of $N \subset \bigcup_{i=1}^\infty B(x_i, r_i)$ such that $\sum_{i=1}^\infty (r_i)^n < \varepsilon$.

  This almost everywhere notation leads to another (easily solvable issue): We sometimes want to say $f = g$ where f is a Lebesgue-integrable function and g is continuous. Clearly a function f which a.e. is equal to a continuous function may not be continuous (same example as above) – so there could be different continuous functions representing an integrable function? No! If $f = g$ a.e. and both f and g are continuous, then $f = g$ *everywhere*. Issue solved.

  So $\|\cdot\|_{L^p(\Omega)}$ is a norm on the collections of functions (where two functions are the same if they are the same a.e.)

  Also: we usually don't care about definining functions everywhere, almost everywhere is enough. Like $\frac{x}{|x|}$ is a function (even though we have no idea what it would be in $x = 0$).
- There is also $L^\infty(\Omega)$, it consists of all (measurable) functions $f : \Omega \to \mathbb{R}$ such that
$$\|f\|_{L^\infty(\Omega)} = \sup_{x\in\Omega} |f(x)| < \infty.$$
Now here is another issue: this supremum cannot be the usual supremum, if not $f = 0$ in $\Omega \setminus \{x_0\}$ and $f(x_0) = 1$ has nonzero-norm (and this is inconsitent with the almost everywhere considerations above). So this is (always, unless otherwise said or forgotten to say) the *essential supremum*:
$$\sup_{x\in\Omega} |f(x)| = \inf \{\Lambda > 0 : \quad f(x) \leq \Lambda \text{ for almost every } x \in \Omega\}$$
Observe it is not true that continuous functions are dense in $L^\infty(\Omega)$. Take $f(x) = 0$ for $|x| < 1$ and $f(x) = 1$ for $|x| > 1$. This is clearly discontinuous, but we see that $\|f\|_{L^\infty(\mathbb{R}^n)} = 1$. If there was a sequence of continuous functions $f_k \in$

C^0, $\|f_k - f\|_{L^\infty(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$ then f_k would converging *uniformly* and the limit of uniformly converging continuous functions must be continuous!

- $L^2(\Omega)$ is a special case since it is a Hilbert space, meaning there is a scalar product

$$\langle f, g\rangle_{L^2} := \int_\Omega fg$$

Now let us define the *Sobolev space* $W^{k,p}(\Omega)$. It means functions with k derivatives in L^p (and it is debated what the W stands for) – often it is handwavingly said $f \in W^{1,p}(\Omega)$ if $f \in L^p(\Omega)$ and any partial derivative $\partial_i f \in L^p(\Omega)$. But how do we take a partial derivative of a (possibly even discontinuous) function f? This leads to the idea of *distributional derivative*. Instead of looking at (e.g. integrable) functions f evaluated at a point x (which by the above discussion we know only works a.e. etc.), we could look at how a function f *acts* on *test-functions* $\varphi \in C_c^\infty(\Omega)$: that is instead of considering

$$f(x) \quad \forall x \in \Omega,$$

why not consider

$$f[\varphi] = \int_\Omega f\,\varphi \quad \forall \varphi \in C_c^\infty(\Omega).$$

This interpreation of a function is called a *distribution* (we will talk about this more when talking about the Fourier transform). The upshot is that this way we can measure every propery of f, just like $f(x)$. The basic lemma for this is the fundamental theorem of Calculus of Variations (for a proof see Lemma 4.45)

Lemma 16.1 (Fundamental lemma of the Calculus of Variations). *Let $f, g \in L^1_{loc}(\Omega)$ (i.e. $f, g \in L^1(\tilde{\Omega})$ whenever $\tilde{\Omega}$ is bounded and $\overline{\tilde{\Omega}} \subset \Omega$) and assume that*

$$f[\varphi] = g[\varphi] \quad \forall \varphi \in C_c^\infty(\Omega)$$

then $f = g$ almost everywhere.

In particular we say that in distributional sense $f = g$ if

$$f[\varphi] = g[\varphi] \quad \forall \varphi \in C_c^\infty(\Omega)$$

And we know that $f = g$ in the integrable sense is the same as $f = g$ in the distributional sense.

Also, if we have a distributional interpretation of a function, then we can take its derivative by a formal integration by parts

$$\partial_i f[\varphi] = \int_\Omega \partial_i f \varphi = -\int_\Omega f \partial_i \varphi = -f[\partial_i \varphi].$$

The middle integral of course makes no sense in general (we would need that f is differentiable), so we just use the definition

$$\partial_i f[\varphi] := -f[\partial_i \varphi].$$

This is called the distributional derivative ∂_i of f. And we say that $\partial_i f = g$ if the two distributions are the same. That is we say $\partial_i f \in L^p(\Omega)$ if there exists $g \in L^p(\Omega)$ (the distributional derivative) if

$$\partial_i f[\varphi] = g[\varphi] \quad \forall \varphi \in C_c^\infty(\Omega),$$

i.e. if

$$-\int f \partial_i \varphi = \int g\, \varphi \quad \forall \varphi \in C_c^\infty(\Omega).$$

If $f \in L^p(\Omega)$ and $\partial_i f \in L^p(\Omega)$ for all $i = 1, \ldots, n$ then we say $f \in W^{1,p}(\Omega)$ (and similarly if $f \in W^{k,p}(\Omega)$.

We can equip $W^{1,p}$ with a norm

$$\|f\|_{W^{1,p}(\Omega)} := \left(\|f\|_{L^p(\Omega)}^p + \sum_{i=1}^n \|\partial_i f\|_{L^p(\Omega)}^p \right)^{\frac{1}{p}}.$$

For $p = 2$, $W^{1,2}(\Omega)$ is still a Hilbert space, with the scalar product

$$\langle f, g \rangle_{W^{1,2}} = \langle f, g \rangle_{L^2} + \sum_{i=1}^n \langle \partial_i f, \partial_i g \rangle_{L^2}$$

One word of warning for distributional derivatives: they are not the same as a.e. pointwise derivatives (which is a concept one should not use): the typical example is the Heaviside function

$$H(x) := \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Formally we might think that its derivative is 0 a.e. – but indeed its *distributional* derivative is not zero (it is the Dirac measure)

$$H'[\varphi] = \varphi(0) \quad \forall \varphi \in C_c^\infty(\Omega).$$

So we should be very careful to compute a derivative almost everywhere and hope thats the distributional derivative (but sometimes this works).

If Ω is a nice set (e.g. a ball, or more generally a set with a non-crazy boundary) then there is another way of thinking of Sobolev functions.

If $p \in [1, \infty)$ we have $f \in W^{1,p}(\Omega)$ if and only if there exists $f_k \in C^\infty(\overline{\Omega}) \cap W^{1,p}(\Omega)$ such that

$$\|f_k - f\|_{W^{1,p}(\Omega)} \xrightarrow{k \to \infty} 0.$$

So again, when talking about Sobolev functions we may assume they are pretty much smooth functions.

What happens when we consider the approximating functions $f_k \in C_c^\infty(\Omega)$? Then we obtain $W_0^{1,p}(\Omega)$, which consists of all maps $f \in W^{1,p}(\Omega)$ such that $f = 0$ on $\partial\Omega$ (in a suitable "trace sense"). This is a good space as well, and it has the nice property that the only constant function in $W_0^{1,p}(\Omega)$ is the constant 0.

Important inequalities are Poincaré inequality, Theorem 13.36, Rellich's theorem, Theorem 13.35, Sobolev inequality, Theorem 13.41, Sobolev embedding, Corollary 13.45, and Morrey embedding Theorem 13.49.

17. Fourier Transform

We follow presentations in [Grafakos, 2014] and [Lieb and Loss, 2001]

17.1. Quick review of complex numbers.

It is way nicer and convenient to consider the complex Fourier transform (even for real functions).

The complex plane $\mathbb{C} \cong \mathbb{R}^2$ consists of points $z = a + \mathbf{i}b$ where $a, b \in \mathbb{R}$ and $\mathbf{i}$ is the imaginary unit that is defined to satisfy $\mathbf{i}^2 = 1$.

For $z = a + \mathbf{i}b$ we call a the real part of z and b the imaginary part of z,

$$\Re\,(a + \mathbf{i}b) := a \quad \Im\,(a + \mathbf{i}b) := b$$

The *complex conjugation* is the map $z \mapsto \overline{z}$ is a map $\mathbb{C} \to \mathbb{C}$ defined as

$$\overline{a + \mathbf{i}b} := a - \mathbf{i}b.$$

We have the norm $|z|$ (which coincides with the $\mathbb{R}^2$-norm)

$$|a + \mathbf{i}b|^2 = a^2 + b^2 = \overline{(a + \mathbf{i}b)}\,(a + \mathbf{i}b)$$

The integral over complex numbers is defined by linearity. I.e. let $f : \mathbb{R}^n \to \mathbb{C}$ then

$$\int_{\mathbb{R}^n} f(x)dx := \int_{\mathbb{R}^n} \Re(f(x))dx + \mathbf{i} \int_{\mathbb{R}^n} \Im(f(x))dx$$

(this is very similar to the integrals of $f : \mathbb{R}^n \to \mathbb{R}^N$, which we also define componentwise).

Observe that for us the *domain* of functions will always be real, but the image space might be $\mathbb{C}$ – i.e. $f : \mathbb{R}^n \to \mathbb{C}$ is the generic function.

17.2. Some motivation.

According to Harmonic Analysts the Fourier transform is the "single most important tool in Harmonic Analysis", [Grafakos, 2014].

Let us illustrate why it is also important for differential equations by doing some well-meant, *formal* calculations.

First we define for $\xi \in \mathbb{R}^n$ the *Fourier transform* of u at ξ

$$\mathcal{F}(f)(\xi) \equiv \hat{f}(\xi) := \int_{\mathbb{R}^n} e^{-2\pi \mathbf{i} \langle \xi, x\rangle} f(x)\, dx.$$

Let us ignore for now

- $\mathcal{F}u(\xi)$ is complex (just use $e^{-2\pi\mathbf{i}\langle\xi,x\rangle} = \cos(-2\pi\langle\xi, x\rangle) + i\sin(-2\pi\langle\xi, x\rangle)$ and integrate imaginary and real part seperately.

- The convergence of the integral might be an issue. It is absolutely convergent if $f \in L^1(\mathbb{R}^n)$, but its unclear what it means e.g. if $f \in L^2(\mathbb{R}^n)$

The inverse Fourier transform

$$\mathcal{F}^{-1}(f)(x) \equiv f^\vee(x) := \int_{\mathbb{R}^n} e^{+2\pi \mathbf{i}\langle \xi, x\rangle} f(\xi)\, d\xi.$$

Assume we want to find a solution to the equation

$$\Delta u(x) = \sum_{i=1}^n \partial_{x^i}\partial_{x^i} u(x) = f(x) \quad \text{in } \mathbb{R}^n \tag{17.1}$$

One of the important properties is that derivatives become polynomial factors after Fourier transform:

$$\left(\partial_{x_i} g\right)^\wedge(\xi) = -i2\pi \mathbf{i}\xi_i \hat{g}(\xi).$$

For the Laplace operator Δ this implies

$$(\Delta u)^\wedge(\xi) = -4\pi^2|\xi|^2\hat{u}(\xi).$$

This means that if we look at the equation (17.1) and apply Fourier transform on both sides we have

$$-|\xi|^2\hat{u}(\xi) = \hat{f}(\xi),$$

that is

$$\hat{u}(\xi) = -|\xi|^{-2}\hat{f}(\xi),$$

Inverting the Fourier transform we get an explicit formula for u in terms of the data f.

$$u(x) = -\left(|\xi|^{-2}\hat{f}(\xi)\right)^\vee(x).$$

This is not a very nice formula, so let us simplify it. Another nice property of Fourier transform (and its inverse) is that products become convolutions. Namely

$$\left(g(\xi)f(\xi)\right)^\vee(x) = \int_{\mathbb{R}^n} g^\vee(x-z) f^\vee(z)\, dz.$$

In our case, for $g(\xi) = -|\xi|^{-2}$ we get that

$$u(x) = \int_{\mathbb{R}^n} g^\vee(x-z)\, f(z)\, dz.$$

Now we need to compute $g^\vee(x-z)$, and for this we restrict our attention to the situation where the dimension is $n \geq 3$. In that case, just by the definition of the (inverse) Fourier transform we can compute that since g has homogeneity of order 2 (i.e. $g(t\xi) = t^{-2}g(\xi)$, then $g^\vee$ is homogeneous of order $2-n$. In particular

$$g^\vee(x) = |x|^{2-n} g^\vee(x/|x|).$$

Now an argument that radial functions stay radial under Fourier transforms (Lemma 17.8 below) leads us to conclude that

$$g^\vee(x) = c_1|x|^{2-n}.$$

That is, we have arrived that (by formal computations) a solution of (17.1) should satisfy

$$u(x) = c_1 \int_{\mathbb{R}^n} |x-z|^{2-n} f(z)\, dz. \tag{17.2}$$

The constant c_1 can be computed explicitely, and one can check that this *potential representation* of a solution u to (17.1) really is true.

17.3. **Precise Definition.** The Fourier transform is a fantastic operator on $L^2(\mathbb{R}^n)$ (and a bit strange operator on $L^1(\mathbb{R}^n)$) however on $L^2(\mathbb{R}^n)$ the integral definition of $\mathcal{F}f$ may not converge absolutely. So we would like to work with nicer dense functions. $C_c^\infty(\mathbb{R}^n)$ comes to mind, but $\mathcal{F}(C_c^\infty(\mathbb{R}^n))$ is not in $C_c^\infty(\mathbb{R}^n)$, this is a version of the *Heisenberg uncertainty principle*.

This is why we will work with *Schwarz classes*

Definition 17.1. A function $f \in C^\infty(\mathbb{R}^n)$ (possibly complex valued) is a *Schwarz function*, in symbols $f \in \mathscr{S}(\mathbb{R}^n)$, if it decays quickly at infinity, more precisely, the *Schwarz seminorms*

$$p_{\alpha,\beta}(f) := \sup_{x\in\mathbb{R}^n} |x|^\alpha |D^\beta f(x)| < \infty \quad \forall \alpha \geq 0, \beta \in \mathbb{N}\cup\{0\}.$$

Exercise 17.2. *Show*

(1) $C_c^\infty(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$
(2) $e^{-|x|^2}$ *is in* $\mathscr{S}(\mathbb{R}^n)$
(3) $e^{-|x|}$ *is not in* $\mathscr{S}(\mathbb{R}^n)$
(4) $\frac{1}{1+|x|^2}$ *is not in* $\mathscr{S}(\mathbb{R}^n)$
(5) the only polynomial in $\mathscr{S}(\mathbb{R}^n)$ *is the constant zero*
(6) if $f \in \mathscr{S}(\mathbb{R}^n)$ *the so is* $\partial^\alpha f$ *for any multiindex* α.
(7) if $f \in \mathscr{S}(\mathbb{R}^n)$ *then so is* $x^\beta f(x)$ *for any multiindex* α.
(8) if $f, g \in \mathscr{S}(\mathbb{R}^n)$ *then the convolution* $f * g \in \mathscr{S}(\mathbb{R}^n)$

Exercise 17.3. *Let* $f \in C^\infty(\mathbb{R}^n)$. *Show that* $f \in \mathscr{S}(\mathbb{R}^n)$ *if and only if for all positive integers* N *and all multi-indices* α *there exists a positive constant* $C_{\alpha,N}$ *such that*

$$|(\partial^\alpha f)(x)| \leq C_{\alpha,N}(1+|x|)^{-N}.$$

Definition 17.4. Let $f \in \mathscr{S}(\mathbb{R}^n)$, then we define the Fourier transform of f as

$$\mathcal{F}(f)(\xi) \equiv \hat{f}(\xi) := \int_{\mathbb{R}^n} e^{-2\pi \mathbf{i}\langle \xi, x\rangle} f(x)\, dx.$$

The inverse Fourier transform of f is given as

$$\mathcal{F}^{-1}(f)(x) \equiv f^\vee(x) := \int_{\mathbb{R}^n} e^{+2\pi \mathbf{i}\langle \xi, x\rangle} f(\xi)\, d\xi.$$

Equivalently, $\mathcal{F}^{-1}(f)(x) = \overline{\mathcal{F}\overline{f}}$.

Observe while we prefer to work with real function $f \in \mathscr{S}(\mathbb{R}^n)$ the Fourier transform is the complex function $\hat{f} : \mathbb{R}^n \to \mathbb{C}$.

Example 17.5. Let $f(x) = e^{-\pi|x|^2}$ then

$$\hat{f}(\xi) = f(\xi).$$

Proof. From Advanced Calculus (this was this Fubini trick!) we have

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}.$$

In particular

$$\int_{-\infty}^{\infty} e^{-\pi t^2} dt = 1.$$

Now consider the map

$$g(s) := \int_{-\infty}^{\infty} e^{-\pi(t+\mathbf{i}s)^2} dt \in \mathbb{C} \cong \mathbb{R}^2.$$

Observe that this integral converges for each s because

$$e^{-\pi(t+\mathbf{i}s)} = e^{-\pi(t^2-s^2)} \underbrace{e^{-2\pi\mathbf{i}st}}_{|\cdot|\leq 1},$$

so

$$|e^{-\pi(t+\mathbf{i}s)}| \leq e^{-\pi(t^2-s^2)}$$

and $t \mapsto e^{-\pi(t^2-s^2)}$ is clearly integrable for any s.

Moroever we have

$$\frac{d}{ds} e^{-\pi(t+\mathbf{i}s)^2} = -2\pi\mathbf{i}(t+\mathbf{i}s)e^{-\pi(t+\mathbf{i}s)^2} = \mathbf{i}\frac{d}{dt} e^{-\pi(t+\mathbf{i}s)^2}$$

Integrating both sides (we can interchange the integration on the left-hand side since we have good convergence of the integral)

$$\frac{d}{ds} g(s) = \mathbf{i} \int_{-\infty}^{\infty} \frac{d}{dt} e^{-\pi(t+\mathbf{i}s)^2} dt = 0.$$

Thus $g'(s) = 0$ for all s, that is $g(s) = g(0)$ for all s. Consequently,

$$\begin{aligned}
\int_{\mathbb{R}} e^{-\pi r^2} e^{-2\pi i t r} dr = \int_{\mathbb{R}} e^{-\pi r^2 - 2\pi i t r} dr = e^{-2\pi t^2} & \int_{\mathbb{R}} e^{-\pi(r^2 - 2itr - 2t^2)} dr \\
= & e^{-2\pi t^2} \underbrace{\int_{\mathbb{R}} e^{-\pi(r+it)^2} dr}_{=g(t)=g(0)=1} \\
= & e^{-2\pi t^2}
\end{aligned}$$

This proves the claim for $n = 1$.

For $n \geq 2$ we use Fubini's theorem

$$
\begin{aligned}
&\int_{\mathbb{R}^n} e^{-\pi|x|^2} e^{-2\pi \mathbf{i}\langle \xi, x\rangle} dx \\
=&\int_{\mathbb{R}} \ldots \left(\int_{\mathbb{R}} e^{-\pi|x_2|^2} e^{-2\pi \mathbf{i}\xi_2 x_2} \left(\int_{\mathbb{R}} e^{-\pi|x_1|^2} e^{-2\pi \mathbf{i}\xi_1 x_1} dx_1\right) dx_2\right) \ldots dx_n \\
=&e^{-\pi|\xi_1|^2} \int_{\mathbb{R}} \ldots \left(\int_{\mathbb{R}} e^{-\pi|x_2|^2} e^{-2\pi \mathbf{i}\xi_2 x_2} dx_2\right) \ldots dx_n \\
=&\ldots \\
=&e^{-\pi|\xi_1|^2} e^{-\pi|\xi_2|^2} \ldots e^{-\pi|\xi_n|^2} \\
=&e^{-\pi|\xi|^2}.
\end{aligned}
$$

$\square$

Exercise 17.6. *Assume that $f \in \mathscr{S}(\mathbb{R}^n)$ and $\mathcal{F}f(\xi) = f(\xi)$ for all $\xi \in \mathbb{R}^n$.*

Show that $f(0) = \int_{\mathbb{R}^n} f(x) dx$

Above we used an important property of the Fourier transform: we can take it componentwise. Let us collect this, and related properties.

Remark 17.7.
- Whenever $f \in \mathscr{S}(\mathbb{R}^\ell), g \in \mathscr{S}(\mathbb{R}^{n-\ell})$ (exercise: show that $fg \in \mathscr{S}(\mathbb{R}^n)$) we have

$$\mathcal{F}_{\mathbb{R}^n}\left(f(x_1,\ldots,x_\ell)g(x_{\ell+1},\ldots,x_n)\right)(\xi) = \mathcal{F}_{\mathbb{R}^\ell}\left(f(x_1,\ldots,x_\ell)\right)(\xi_1,\ldots,\xi_\ell)\mathcal{F}_{\mathbb{R}^{n-\ell}}\left(g(x_{\ell+1},\ldots,x_n)\right)(\xi_{\ell+1},\ldots,\xi_n).$$

  This follows, as above, from Fubini's theorem.

The Fourier transform has many nice properties (most importantly how it behaves with respect to differentiation!)

Lemma 17.8. *Let $f, g \in \mathscr{S}(\mathbb{R}^n)$. Then*

(1) $\|\hat{f}\|_{L^\infty(\mathbb{R}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}$.
(2) $\widehat{f+g} = \hat{f} + \hat{g}$
(3) $\widehat{\lambda f} = \lambda \hat{f}$ *for any complex* $\lambda \in \mathbb{C}$
(4) Let
$$\tilde{f}(x) := f(-x).$$
Then
$$\hat{\tilde{f}} = \tilde{\hat{f}}$$
(5) Fix $y \in \mathbb{R}^n$. Denote by $\tau_y : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ the translation by y, i.e.
$$\tau_y f(x) = f(x+y).$$
Then
$$\mathcal{F}(\tau_y f)(\xi) = e^{2\pi \mathbf{i}\langle \xi, y\rangle} \mathcal{F}f(\xi)$$

(6) *Denote the scaling* $\delta_t f(x) := f(tx)$ *for some fixed* $t > 0$. *Then*

$$\mathcal{F}(\delta_t f)(\xi) = t^{-n}\delta_t(\mathcal{F}(\xi)).$$

(7) $\mathcal{F}(\partial_{x^k} f)(\xi) = 2\pi \mathbf{i} \xi^k \mathcal{F}(\xi)$

(8) $\partial_{\xi^k}\mathcal{F}(\xi) = \mathcal{F}\left(-2\pi\mathbf{i}x^k f(x)\right)(\xi)$

(9) *If* $P \in O(n)$, *i.e.* $P^t P = I$ *then* $\mathcal{F}(f(P\cdot))(\xi) = \mathcal{F}f(P\xi)$.

(10) *Let* $f \in \mathscr{S}(\mathbb{R}^n, \mathbb{C})$ *then* $\overline{\mathcal{F}f(\xi)} = \mathcal{F}\overline{f}(-\xi)$

(11) *If* f *is real and even,* $f(x) = f(-x)$ *then* $\mathcal{F}f$ *is real number*

(12) *If* f *is real and odd,* $f(x) = -f(-x)$ *then* $\mathcal{F}f$ *is imaginary*

(13) *If* f *is radial, i.e. if* f *can be written as* $f(x) = g(|x|)$, *then* $\mathcal{F}f$ *is radial.*

(14) *If* f *is positively* σ*-homogeneous for some* $\sigma \in \mathbb{R}$, *i.e.* $f(\lambda x) = \lambda^\sigma$, *then* $\mathcal{F}f$ *is* $-n-\sigma$*-homogeneous, i.e.* $\mathcal{F}f(\lambda\xi) = \lambda^{-n-\sigma}\mathcal{F}f(\xi)$

(15) *If* $f, g \in \mathscr{S}(\mathbb{R}^n)$ *then* $\mathcal{F}(f * g)(\xi) = \hat{f}(\xi)\hat{g}(\xi)$

Proof. (1) We have

$$|\mathcal{F}f(\xi)| = \left|\int_{\mathbb{R}^n} e^{-2\pi\mathbf{i}\langle\xi,x\rangle} f(x)dx\right| \leq \int_{\mathbb{R}^n} \underbrace{\left|e^{-2\pi\mathbf{i}\langle\xi,x\rangle}\right|}_{=1} |f(x)|\, dx$$

(2) exercise

(3) exercise

(4) exercise

(5) We have

$$\begin{aligned}
\mathcal{F}(\tau_y f)(\xi) &= \int_{\mathbb{R}^n} f(x+y)e^{-2\pi\mathbf{i}\langle\xi,x\rangle}dx \\
&= e^{-2\pi\mathbf{i}\langle\xi,-y\rangle}\int_{\mathbb{R}^n} f(x+y)e^{-2\pi\mathbf{i}\langle\xi,x+y\rangle}dx \\
&= e^{-2\pi\mathbf{i}\langle\xi,-y\rangle}\int_{\mathbb{R}^n} f(z)e^{-2\pi\mathbf{i}\langle\xi,z\rangle}dz \\
&= e^{-2\pi\mathbf{i}\langle\xi,-y\rangle}\mathcal{F}f(\xi).
\end{aligned}$$

(6) exercise

(7) We have using integration by parts (recall that $\lim_{|x|\to\infty} f(x) = 0$)

$$\begin{aligned}
\mathcal{F}(\partial_{x^k} f)(\xi) &= \int_{\mathbb{R}^n} \partial_{x^k} f(x)e^{-2\pi\mathbf{i}\langle\xi,x\rangle}dx \\
&= -\int_{\mathbb{R}^n} f(x)\partial_{x^k}\left(e^{-2\pi\mathbf{i}\langle\xi,x\rangle}\right)dx \\
&= -\left(-2\pi\mathbf{i}\xi^k\right)\int_{\mathbb{R}^n} f(x)e^{-2\pi\mathbf{i}\langle\xi,x\rangle}dx \\
&= 2\pi\mathbf{i}\xi^k\mathcal{F}f(\xi).
\end{aligned}$$

(8) exercise

(9) We will use the transformation rule for $\Phi(x) := Px$, which satisfied $\det(D\Phi) = \det(P) = 1$. Also observe that $\langle \xi, x\rangle = \langle P\xi, Px\rangle$ if P is orthogonal.

$$\begin{aligned}\mathcal{F}(f(P\cdot))(\xi) =& \int_{\mathbb{R}^n} f(Px)e^{-2\pi \mathbf{i}\langle \xi,x\rangle}dx\\ =& \int_{\mathbb{R}^n} f(Px)e^{-2\pi \mathbf{i}\langle P\xi,Px\rangle}dx\\ =& \int_{\mathbb{R}^n} f(z)e^{-2\pi \mathbf{i}\langle P\xi,z\rangle}dx\\ =&\mathcal{F}(f)(P\xi).\end{aligned}$$

(10) We have

$$\begin{aligned}\overline{\mathcal{F}(f(\cdot))(\xi)} =& \int_{\mathbb{R}^n} \overline{f(x)e^{-2\pi \mathbf{i}\langle \xi,x\rangle}}dx\\ =& \int_{\mathbb{R}^n} \overline{f(x)}\,\overline{e^{-2\pi \mathbf{i}\langle \xi,x\rangle}}dx\\ =& \int_{\mathbb{R}^n} \overline{f(x)}e^{+2\pi \mathbf{i}\langle \xi,x\rangle}dx\\ =& \int_{\mathbb{R}^n} \overline{f(x)}e^{-2\pi \mathbf{i}\langle -\xi,x\rangle}dx\\ =& \int_{\mathbb{R}^n} \overline{f(x)}e^{-2\pi \mathbf{i}\langle -\xi,x\rangle}dx\end{aligned}$$

(11) exercise
(12) exercise
(13) f is radial if and only if there exist g such that $f(x) = g(|x|)$ for all $x \in \mathbb{R}^n$. This is equivalent to saying $f(Px) = f(x)$ for all $x \in \mathbb{R}^n$ and $P \in O(n)$. Indeed set $g(r) := f(re)$ for $r \in \mathbb{R}$ and $e \in \mathbb{R}^n$ any unit vector (i.e. $|e| = 1$). For $x \neq 0$ denote by $P \in O(n)$ the rotation such that $Px = |x|e$. Then $f(x) = f(P|x|e) = f(|x|e) = g(|x|)$.

Now if f is radial we have for any fixed $P \in O(n)$ from the observations above

$$\mathcal{F}(f)(\xi) = \mathcal{F}(f(P\cdot))(\xi) = \mathcal{F}f(P\xi).$$

So $\mathcal{F}(f)$ is radial.
(14) exercise

(15) We have (using Fubini's theorem since everything converges)

$$\begin{aligned}
\mathcal{F}(f*g)(\xi) &= \int_{\mathbb{R}^n} (f*g)(x) e^{-2\pi \mathbf{i}\langle x,\xi\rangle} dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)g(y) dy e^{-2\pi \mathbf{i}\langle x,\xi\rangle} dx \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x-y)g(y) e^{-2\pi \mathbf{i}\langle x-y,\xi\rangle} e^{-2\pi \mathbf{i}\langle y,\xi\rangle} dxdy \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(x-y) e^{-2\pi \mathbf{i}\langle x-y,\xi\rangle} dx \right) g(y) e^{-2\pi \mathbf{i}\langle y,\xi\rangle} dy \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f(z) e^{-2\pi \mathbf{i}\langle z,\xi\rangle} dz \right) g(y) e^{-2\pi \mathbf{i}\langle y,\xi\rangle} dy \\
&= \int_{\mathbb{R}^n} (\mathcal{F}f(\xi))\, g(y) e^{-2\pi \mathbf{i}\langle y,\xi\rangle} dy \\
&= (\mathcal{F}f(\xi)) \int_{\mathbb{R}^n} g(y) e^{-2\pi \mathbf{i}\langle y,\xi\rangle} dy \\
&= (\mathcal{F}f(\xi))\,(\mathcal{F}g(\xi)).
\end{aligned}$$

□

Exercise 17.9. *Proof the remaining statements of Lemma 17.8.*

Here is the main reason we prefer to work with $\mathscr{S}(\mathbb{R}^n)$ when dealing with the Fourier transform: It maps $\mathscr{S}(\mathbb{R}^n)$ to $\mathscr{S}(\mathbb{R}^n)$ (but as we shall see below it does not map $C_c^\infty(\mathbb{R}^n)$ to $C_c^\infty(\mathbb{R}^n)$, Theorem 17.14).

Theorem 17.10. *Let $f \in \mathscr{S}(\mathbb{R}^n)$ then $\hat{f} \in \mathscr{S}(\mathbb{R}^n)$*

Proof. We have

$$D^\beta \hat{f}(x) \leq c\widehat{x^\beta f(x)}(\xi).$$

Since $f \in \mathscr{S}(\mathbb{R}^n)$ we have $g(x) = x^\beta f(x) \in \mathscr{S}(\mathbb{R}^n)$ (see Exercise 17.2).

In particular we have by Lemma 17.8

$$\|D^\beta \hat{f}(x)\|_{L^\infty} \leq \||\cdot|^{|\beta|} f(\cdot)\|_{L^1} < \infty,$$

that is $\hat{f}$ is $C^\infty(\mathbb{R}^n)$ since f, Df, etc. are all Lipschitz.

So we need to show that for $g \in \mathscr{S}(\mathbb{R}^n)$ we have for any $K = 0, 1, 2, \ldots,$.

$$\sup_{\xi \in \mathbb{R}^n} |\xi|^{2K} |\hat{g}(\xi)| < \infty$$

For $K = 0$ we have by Lemma 17.8

$$\|\hat{g}\|_{L^\infty(\mathbb{R}^n)} \leq \|g\|_{L^1(\mathbb{R}^n)}$$

For $K = 1$ observe that

$$|\xi|^2 \hat{h}(\xi) = (-\Delta h)^\wedge(\xi),$$

where $\Delta = \sum_\alpha \partial_\alpha \partial_\alpha$. Since $g \in \mathscr{S}(\mathbb{R}^n)$ so is Δg.

So if we set $g_K := \Delta g_{K-1}$ (with $g_0 = g$) we have that $g_K \in \mathscr{S}(\mathbb{R}^n)$ and

$$\sup_\xi |\xi|^{2K} |\hat{g}(\xi)| = c \sup_\xi |\hat{g_K}(\xi)| < \infty.$$

□

We will show below in Theorem 17.14 that the same is not true for $C_c^\infty(\mathbb{R}^n)$, i.e. $\mathcal{F} C_c^\infty(\mathbb{R}^n) \not\subset C_c^\infty(\mathbb{R}^n)$.

But for this we need the following fundamental theorem on the *Fourier inverse*. Recall that the (*complex*) L^2-scalar product is

$$\langle f, g \rangle = \int_{\mathbb{R}^n} f \overline{g} d\mathcal{L}^n.$$

Observe that in this sense

$$\mathcal{F} f(\xi) = \langle f, e^{2\pi \mathbf{i} \langle \xi, \cdot \rangle} \rangle.$$

Now in *some* sense (none of the integrals makes really sense)

$$\langle e^{2\pi \mathbf{i} \langle \xi, \cdot \rangle}, e^{2\pi \mathbf{i} \langle \eta, \cdot \rangle} \rangle = \delta_{\xi,\eta} = \begin{cases} \infty & \xi = \eta \\ 0 & \xi \neq \eta. \end{cases}$$

I.e. $\xi \mapsto e^{-2\pi \mathbf{i} \langle \xi, \cdot \rangle}$ is an orthonormal system in $L^2(\mathbb{R}^n)$. So in *some* sense we have by *formal* Fubini's theorem

$$\begin{aligned}
\mathcal{F}^{-1}(\mathcal{F} f)(x) = &\int_{\mathbb{R}^n} e^{2\pi \mathbf{i} \langle x, \xi \rangle} \langle f, e^{2\pi \mathbf{i} \langle \xi, \cdot \rangle} \rangle d\xi = \int_{\mathbb{R}^n} f(y) \langle e^{2\pi \mathbf{i} \langle x, \cdot \rangle}, e^{2\pi \mathbf{i} \langle \cdot, y \rangle} \rangle dy \\
= &\int_{\mathbb{R}^n} f(y) \delta_{y,x} dy \\
= &f(x).
\end{aligned}$$

The fun part about this computation is that it actually works. Which leads to the most important fundamental theorem of Fourier Analysis.

Theorem 17.11 (Fourier inverse)**.** *Let $f, g \in \mathscr{S}(\mathbb{R}^n)$ then the following is true.*

(1)

$$\int_{\mathbb{R}^n} f(x) \hat{g}(x) dx = \int_{\mathbb{R}^n} \hat{f}(x) g(x) dx$$

(2) Fourier transform inversion

$$\mathcal{F}^{-1}(\mathcal{F} f)(x) = \mathcal{F}\left(\mathcal{F}^{-1} f\right)(x) = f(x) \quad \forall x \in \mathbb{R}^n$$

(3) Parseval's relation

$$\int_{\mathbb{R}^n} f(x) \overline{g(x)} = \int_{\mathbb{R}^n} \hat{f}(\xi) \overline{\hat{g}(\xi)} d\xi$$

or equivalently,

$$\langle f, g\rangle = \langle \hat{f}, \hat{g}\rangle.$$

(4)

$$\int_{\mathbb{R}^n} f(x)g(x)dx = \int_{\mathbb{R}^n} \hat{f}(\xi)\, g^{\vee}(\xi)d\xi.$$

(5) Plancherell identity

$$\|f\|_{L^2(\mathbb{R}^n)} = \|\hat{f}\|_{L^2(\mathbb{R}^n)} = \|f^{\vee}\|_{L^2(\mathbb{R}^n)}$$

Proof. (1) We have (observe all integrals converge, so we can use Fubini's theorem

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\,\hat{g}(x)dx &= \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^n} e^{-2\pi\mathbf{i}\langle x,y\rangle} g(y)dy\,dx\\ &= \int_{\mathbb{R}^n} g(y) \int_{\mathbb{R}^n} f(x)e^{-2\pi\mathbf{i}\langle x,y\rangle} dx\,dy\\ &= \int_{\mathbb{R}^n} g(y)\,\hat{f}(y)\,dy.\end{aligned}$$

(2) Fix $t \in \mathbb{R}^n$ and $\varepsilon > 0$ and set

$$g(\xi) := e^{2\pi i\langle \xi,t\rangle}e^{-\pi|\varepsilon\xi|^2}$$

Observe this is the same we did in the formal computation before the theorem, essentially $g(\xi) = e^{2\pi i\langle\xi,t\rangle}$ times the dampening $e^{-\pi|\varepsilon\xi|^2}$ which ensures that $h \in \mathscr{S}(\mathbb{R}^n)$ for any $\varepsilon > 0$.

By Lemma 17.8 we can compute the Fourier transform: the factor $e^{2\pi i\langle\xi,t\rangle}$ becomes a translation; the Fourier transform of $e^{-\pi|\varepsilon\xi|^2}$ is up to a scaling computed in Example 17.5. So we find

$$\hat{g}(x) = \frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}$$

Observe that $\hat{g}$ is an approximation of the identity (i.e. as $\varepsilon \to 0$ it goes to zero for every $x \neq t$, and to ∞ at $x = t$).

From (1) we have

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx &= \int_{\mathbb{R}^n} f(x)\hat{g}(x)dx = \int_{\mathbb{R}^n} \hat{f}(x)g(x)dx\\ &= \int_{\mathbb{R}^n} \hat{f}(x)e^{2\pi i\langle x,t\rangle}e^{-\pi|\varepsilon x|^2}dx\end{aligned} \tag{17.3}$$

Now since $f, \hat{f} \in \mathscr{S}(\mathbb{R}^n)$ both integrals make sense. Moreover, by Lebesgue dominated convergence theorem

$$\lim_{\varepsilon\to 0}\int_{\mathbb{R}^n} \hat{f}(x)e^{2\pi i\langle x,t\rangle}e^{-\pi|\varepsilon x|^2}dx = \int_{\mathbb{R}^n} \hat{f}(x)e^{2\pi i\langle x,t\rangle}dx = \mathcal{F}^{-1}(\hat{f})(t) \tag{17.4}$$

We also have in view of Exercise 17.6 (and substituting $z = \frac{x-t}{\varepsilon}$)

$$\int_{\mathbb{R}^n} \frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx = \int_{\mathbb{R}^n} e^{-\pi|z|^2}dz = e^{-\pi|0|^2} = 1.$$

So
$$\int_{\mathbb{R}^n} f(x)\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx = f(t) + \int_{\mathbb{R}^n} (f(x)-f(t))\,\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx$$
and thus
$$\left|\int_{\mathbb{R}^n} f(x)\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx - f(t)\right| \le \int_{\mathbb{R}^n} |f(x)-f(t)|\,\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx$$
Since $f \in \mathscr{S}(\mathbb{R}^n)$ in particular f is globally Lipschitz, i.e. for some $L \in \mathbb{R}$ we have $|f(x)-f(t)| \le L|x-t|$ for any $x,t \in \mathbb{R}^n$ (we can take $L := \|\nabla f\|_{L^\infty}$). Consequently,

(17.5)
$$\begin{aligned}\left|\int_{\mathbb{R}^n} f(x)\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx - f(t)\right| &\le L\int_{\mathbb{R}^n} |x-t|\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx\\ &= L\varepsilon\int_{\mathbb{R}^n} \left|\frac{x-t}{\varepsilon}\right|\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx\\ &= L\varepsilon\underbrace{\int_{\mathbb{R}^n} |z|\,e^{-\pi|z|^2}dz}_{<\infty}\\ &\xrightarrow{\varepsilon\to 0} 0.\end{aligned}$$

Thus, letting $\varepsilon \to 0$ in (17.3), using (17.4) and (17.5) we conclude
$$f(t) \xleftarrow{\varepsilon\to 0} \int_{\mathbb{R}^n} f(x)\frac{1}{\varepsilon^n}e^{-\pi\left|\frac{x-t}{\varepsilon}\right|^2}dx = \int_{\mathbb{R}^n} \hat{f}(x)e^{2\pi i\langle x,t\rangle}e^{-\pi|\varepsilon x|^2}dx \xrightarrow{\varepsilon\to 0} \mathcal{F}^{-1}(\hat{f})(t).$$
The other equation in (2) follows similarly.

(3) With the help of (2) and (1) we have
$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\overline{g(x)}dx &= \int_{\mathbb{R}^n} \mathcal{F}^{-1}(\mathcal{F}f)(x)\overline{g(x)}dx\\ &= \int_{\mathbb{R}^n} (\mathcal{F}f)(x)\mathcal{F}^{-1}\overline{g(x)}dx\\ &= \int_{\mathbb{R}^n} (\mathcal{F}f)(x)\overline{\mathcal{F}g(x)}dx\end{aligned}$$

(4) exercise (cf. (3))

(5) We have with (3)
$$\|f\|^2_{L^2(\mathbb{R}^n)} = \langle f,f\rangle = \langle \hat{f},\hat{f}\rangle = \|\hat{f}\|^2_{L^2(\mathbb{R}^n)}.$$

□

The following looks ridiculously complicated, but is very useful for pseudo-differential operators (which are defined by Fourier transform)

Example 17.12 (A ridiculous way to integrate by parts)**.** Let $f,g \in \mathscr{S}(\mathbb{R}^n)$. Then for $i = 1,\ldots,n$
$$\int f\partial_i g = -\int \partial_i f\, g$$

Indeed, we have from Parseval, Theorem 17.11,

$$\begin{aligned}\int f\partial_i g &= \int \hat{f}(\xi)\,(\partial_i g)^{\vee}\,(\xi)\\ &= \int \hat{f}(\xi)(+2\pi\mathbf{i}\xi^i)g^{\vee}(\xi)(\xi)\\ &= \int (+2\pi\mathbf{i}\xi^i)\hat{f}(\xi)g^{\vee}(\xi)(\xi)\\ &= -\int(-2\pi\mathbf{i}\xi^i)\hat{f}(\xi)g^{\vee}(\xi)(\xi)\\ &= -\int \widehat{\partial_i f}(\xi)g^{\vee}(\xi)(\xi)\\ &= -\int \partial_i f g.\end{aligned}$$

Exercise 17.13. *Show: if $f,g\in\mathscr{S}(\mathbb{R}^n)$ then $\mathcal{F}(fg)(\xi)=\hat{f}*\hat{g}(\xi)$.*

Hint: Use Theorem 17.11 and Lemma 17.8.

So why do we work with $\mathscr{S}(\mathbb{R}^n)$ and not $C_c^\infty(\mathbb{R}^n)$? Because the Fourier transform does not take $C_c^\infty(\mathbb{R}^n)$ to $C_c^\infty(\mathbb{R}^n)$ indeed (this is a version of the Heisenberg uncertainty principle: you cannot localize in physical space *and* phase space at the same time)

Theorem 17.14. *Let $f\in C_c^\infty(\mathbb{R}^n)$ and assume that $\hat{f}\in C_c^\infty(\mathbb{R}^n)$. Then $f\equiv 0$.*

Proof. The proof uses a bit of complex analysis. First if $f,\hat{f}\in C_c^\infty(\mathbb{R}^n)$ then so is the restriction to $\mathbb{R}\times\{c_2,\ldots,c_n\}$ for any vector $c=(c_2,\ldots,c_n)\in\mathbb{R}^{n-1}$. If we show that $f(\cdot,c_2,\ldots,c_n)$ is zero for every such vector c then f is zero everywhere. So w.l.o.g. $n=1$.

Define F as

$$F(\xi+\mathbf{i}\eta):=\int_{\mathbb{R}} e^{2\pi\mathbf{i}\,\xi x}e^{-2\pi\eta x}f(x)dx$$

F is called *Fourier–Laplace transform* of f, it is an extension of $\mathcal{F}f:\mathbb{R}\to\mathbb{C}$ to $F:\mathbb{C}\to\mathbb{C}$.

F has good decay properties (made precise by the so-called *Paley-Wiener theorem* or in this case *Schwarz's theorem*), for us it is sufficient to observe that $F\in C^1(\mathbb{R}^2)$.

Indeed, *since $f\in C_c^\infty(\mathbb{R}^n)$* there exists some $R>0$ such that

$$x\mapsto e^{2\pi\mathbf{i}\,\xi x}e^{-2\pi\eta x}f(x)\in C_c^\infty(-R,R),$$

and

$$\begin{aligned}\partial_\xi e^{2\pi\mathbf{i}\,\xi x}e^{-2\pi\eta x}f(x) &= 2\pi\mathbf{i}x\,e^{2\pi\mathbf{i}\,\xi x}e^{-2\pi\eta x}f(x),\\ \partial_\eta e^{2\pi\mathbf{i}\,\xi x}e^{-2\pi\eta x}f(x) &= -2\pi\,x\,e^{2\pi\mathbf{i}\,\xi x}e^{-2\pi\eta x}f(x).\end{aligned}$$

Then either using the differentiation rules for the Riemann integral (since everything is continuous and compactly supported, Lebesgue and Riemann integral coincide) or writing down the difference quotient and using the dominated convergence theorem, we conclude $F \in C^1(\mathbb{R}^2)$ and

$$|F(\xi + \mathbf{i}\eta)| \le \|f\|_{L^1(\mathbb{R}^n)} e^{\theta_N |\eta|}$$
$$|\partial_\xi F(\xi + \mathbf{i}\eta)| \le 2\pi \||x| f\|_{L^1(\mathbb{R}^n)} e^{\theta_N |\eta|}$$
$$|\partial_\eta F(\xi + \mathbf{i}\eta)| \le 2\pi \||x| f\|_{L^1(\mathbb{R}^n)} e^{\theta_N |\eta|}.$$

where we choose θ_N so large that $f(x) \equiv 0$ whenever $|x| \ge \frac{\theta_N}{2\pi}$.

Moreover F is holomorphic as a function from $\mathbb{C} \to \mathbb{C}$. Namely one can check the Cauchy-Riemann equations holds:

$$\partial_\xi \Re(F) = \partial_\eta \Im(F)$$

and

$$\partial_\eta \Re(F) = -\partial_\xi \Im(F).$$

Alternatively one writes

$$e^{2\pi \mathbf{i}\,\xi x} e^{-2\pi \eta x} f(x) = e^{2\pi \mathbf{i}\,(\xi + \mathbf{i}\eta) x} f(x)$$

and uses that $z \mapsto e^{2\pi \mathbf{i} z}$ is holomorphic. Holomorphic functions are analytic (i.e. they are locally power functions).

But now since F is analytic, $\hat{f}(\xi) = F(\xi + \mathbf{i}0)$ can written as a power function around any $\xi_0 \in \mathbb{R}$, and for some positive radius of convergence $r > 0$ and $(a_k) \in \mathbb{C}$

$$\hat{f}(\xi) = \sum_{k=0}^{\infty} a_k (\xi - \xi_0)^k \quad \forall |\xi - \xi_0| < r.$$

But now one can show (exercise!) that this implies that if $\hat{f}(\xi_0) = 0$ then either $\hat{f}(\xi) \neq 0$ for all $\xi \approx \xi_0$ or otherwise $\hat{f} \equiv 0$.

Since $f \in C_c^\infty(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ we have with Theorem 17.11 $f(x) = \mathcal{F}^{-1}\hat{f}(x) = \mathcal{F}^{-1}0 = 0$ for all $x \in \mathbb{R}$. □

17.4. **Tempered Distributions and their Fourier transform.** We want to use Fourier transforms for way more general functions. Plancherell Theorem 17.11 gives us an easy way to compute the Fourier transform for L^2, but what about functions like $f(x) = |x|^{-\alpha}$ (see Theorem 17.33)?.

We do a trick similar to the weak (distributional derivatives) and argue via test-functions. But we need to know what convergence is in the space of test functions, so let us be precise:

Definition 17.15 (test functions)**.** We consider two types of testfunctions, $C_c^\infty(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$.

(1) We set $\mathcal{D}(\mathbb{R}^n) := C_c^\infty(\mathbb{R}^n)$ and say that a sequence $f_k \in \mathcal{D}(\mathbb{R}^n)$ converges in $\mathcal{D}(\mathbb{R}^n)$ to some $f : \mathbb{R}^n \to \mathbb{R}$
   - there exists a compact set K such that $\operatorname{supp} f_k \subset K$ for all $k \in \mathbb{N}$ and

- for any multiindex α we have $\|\partial^\alpha(f_k - f)\|_{L^\infty(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0$

It is easy to see that then $f \in \mathcal{D}(\mathbb{R}^n)$ (exercise!)

(2) We say that a sequence $f_k \in \mathscr{S}(\mathbb{R}^n)$ converges in $\mathscr{S}(\mathbb{R}^n)$ to some $f : \mathbb{R}^n \to \mathbb{R}$ if for any multiindex α, β we have

$$\|x^\beta \partial^\alpha(f_k - f)\|_{L^\infty(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

It is easy to see that then $f \in \mathscr{S}(\mathbb{R}^n)$ (exercise!).

Exercise 17.16. *Show the following: let $f \in \mathscr{S}(\mathbb{R}^n)$ then there exist $f_k \in C_c^\infty(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$ with f_k converges to f in the sense of $\mathscr{S}(\mathbb{R}^n)$.*

Exercise 17.17. *Show that the Fourier transform $\mathcal{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is continuous. I.e. assume that $f_k \in \mathscr{S}(\mathbb{R}^n)$ converges to $f \in \mathscr{S}(\mathbb{R}^n)$ in the sense of $\mathscr{S}(\mathbb{R}^n)$. Show that $\mathcal{F} f_k$ also converges to $\mathcal{F} f$ in the sense of $\mathscr{S}(\mathbb{R}^n)$.*

Distributions are dual spaces of $\mathcal{D}$ or $\mathscr{S}$.

Definition 17.18 (Distributions and tempered distributions)**.**

- $\mathcal{D}'(\mathbb{R}^n)$ is the space of continuous linear functionals on $D(\mathbb{R}^n) = C_c^\infty(\mathbb{R}^n)$. More precisely, a linear operator

$$L : C_c^\infty(\mathbb{R}^n) \to \mathbb{R}$$

belongs to $\mathcal{D}'(\mathbb{R}^n)$ if and only if for any $f_k \in C_c^\infty(\mathbb{R}^n)$ converging to f in $\mathcal{D}$ then

$$Lf_k \xrightarrow{k\to\infty} Lf \quad \text{in } \mathbb{R}.$$

$\mathcal{D}'$ is a linear space, called the space of *distributions on $\mathbb{R}^n$*.

- $\mathcal{S}'(\mathbb{R}^n)$ (space of *tempered distribution*s is a subspace of $\mathcal{D}'(\mathbb{R}^n)$ (space of *distributions*). A linear functional L on $\mathscr{S}(\mathbb{R}^n)$ (and thus on $\mathcal{D}(\mathbb{R}^n)$) belongs to $\mathscr{S}'(\mathbb{R}^n)$ if and only if for any $f_k \in \mathscr{S}(\mathbb{R}^n)$ converging to f in $\mathscr{S}(\mathbb{R}^n)$ then

$$Lf_k \xrightarrow{k\to\infty} Lf \quad \text{in } \mathbb{R}.$$

- $f \in \mathcal{D}'$ belongs to L^p if there exists $g \in L^p(\mathbb{R}^n)$ such that

$$f[\varphi] = \int g\varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$$

- We identify functions $f : \mathbb{R}^n \to \mathbb{R}$ with

$$f[\varphi] := \int_{\mathbb{R}^n} f\varphi dx$$

and for measures μ

$$\mu[\varphi] := \int_{\mathbb{R}^n} f d\mu$$

Observe these objects may or may not define distributions or tempered distributions.

- A distribution f is equal to g if $f[\varphi] = g[\varphi]$ for all $\varphi \in C_c^\infty(\mathbb{R}^n)$. (for tempered distributions cf. Exercise 17.19)

- Derivative of a distribution is always defined: We simply set

$$\partial_\alpha f[\varphi] := -f[\partial_\alpha \varphi]$$

Exercise 17.19. *Let $f, g \in \mathscr{S}'(\mathbb{R}^n)$. Show that the following are equivalent.*

- $f[\varphi] = g[\varphi]$ *for all* $\varphi \in C_c^\infty(\mathbb{R}^n)$
- $f[\varphi] = g[\varphi]$ *for all* $\varphi \in \mathscr{S}(\mathbb{R}^n)$

Hint: Exercise 17.16.

Example 17.20. We said when talking about Sobolev spaces,

$$\partial_\alpha f = g$$

if

$$\int f \partial_\alpha \varphi = -\int g\varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$$

So this is equivalent to saying that $\partial_\alpha f$ (as a (tempered) distribution) is equal to g

Proposition 17.21. *$f \in \mathcal{D}'$ belongs to $L^p(\mathbb{R}^n)$, $1 < p < \infty$ if and only if*

$$|f[\varphi]| \leq \Lambda \|\varphi\|_{L^{p'}} \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n)$$

where $p' = \frac{p}{p-1}$.

That is we have

$$f[\varphi] = \int_{\mathbb{R}^n} fg$$

where $g \in L^p(\mathbb{R}^n)$ and moreover $\|g\|_{L^p(\mathbb{R}^n)} \leq \Lambda$.

Proof. This is a consequence of Riesz representation theorem, Exercise 5.42. □

However, many distributions are not functions.

Exercise 17.22.

- *Let f be defined as follows (for some fixed $\ell \in \{1, \ldots, n\}$)*

$$f[\varphi] := \int_{\mathbb{R}^n} \partial_\alpha \varphi dx$$

Show that f is a tempered distribution, i.e. $f \in \mathscr{S}'$.

- *Let μ be a Radon measure, and $\mu(\mathbb{R}^n) < \infty$ and set*

$$f[\varphi] := \int_{\mathbb{R}^n} f d\mu$$

Show that f is a tempered distribution.

- *Let*

$$f[\varphi] := \varphi(0)$$

Show that f is a tempered distribution.

What is $\partial_{x^\alpha} f$?

- *Let*
$$f[\varphi] := \int_{\mathbb{R}^n} e^{|x|^2}\varphi(x)dx$$
Show that f is a distribution, but not a tempered distribution.
- *Give an example of a Radon measure μ on $\mathbb{R}^n$ so that μ is no tempered distribution. Hint: What L^1_{loc}-function is not a tempered distribution*
- *Let $f \in L^1_{loc}(\mathbb{R}^n)$. Show that f is a distribution , but may not be a tempered distribution.*
- *Let $f \in L^p(\mathbb{R}^n)$, $p \in [1,\infty]$, show that*
$$f[\varphi] = \int_{\mathbb{R}^n} f(x)\varphi(x)$$
is a tempered distribution.

So in general we can identify functions (and measures) with their distributions (if they are). If we define the Fourier transform on distributions, we have a huge class of functions whose Fourier transform is defined. To do so, we use Theorem 17.11(1) (which is similar to the integration by parts we use for weak derivatives)

Definition 17.23 (Fourier transform on tempered distributions)**.** Let $f \in \mathscr{S}'(\mathbb{R}^n)$. Then $\hat{f} = \mathcal{F}f \in \mathscr{S}(\mathbb{R}^n)$ is defined as
$$\hat{f}[\varphi] := f[\hat{\varphi}]$$
Observe that $\mathcal{F}f$ is a tempered distribution by Exercise 17.17. Thus $\mathcal{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is linear.

Example 17.24.
- Fix x_0 and let $\delta_{x_0}[\varphi] := \varphi(x_0)$ which is clearly a tempered distribution. Then
$$\widehat{\delta_{x_0}}[\varphi] = \hat{\varphi}(x_0) = \int_{\mathbb{R}^n} \varphi(x)e^{-2\pi \mathbf{i}\langle x,x_0\rangle}\, dx$$
So, $\widehat{\delta_{x_0}} = e^{-2\pi\mathbf{i}\langle x,x_0\rangle}$.
In particular $\widehat{\delta_0} = 1$
- Conversely the Fourier transform of the constant 1 is the dirac delta at zero, $\mathcal{F}[1] = \delta_0$. Indeed, $f(x) \equiv 1$ is a tempered distribution and
$$\mathcal{F}1[\varphi] = 1[\mathcal{F}\varphi] = \int_{\mathbb{R}^n} \mathcal{F}\varphi(x)dx = \mathcal{F}^{-1}\left(\mathcal{F}\varphi\right)(0) = \varphi(0) = \int_{\mathbb{R}^n} \varphi d\delta_0.$$
- Let $p(x) = x_k$. Then in view of Lemma 17.8,
$$\begin{aligned}\hat{p}[\varphi] = &\int_{\mathbb{R}^n} x_k\mathcal{F}\varphi(x)dx = \frac{1}{2\pi\mathbf{i}}\int_{\mathbb{R}^n} 2\pi\mathbf{i}x_k\mathcal{F}\varphi(x)dx = \frac{1}{2\pi\mathbf{i}}\int_{\mathbb{R}^n} \mathcal{F}\left(\partial_{x_k}\varphi\right)(x)dx\\ =&\frac{1}{2\pi\mathbf{i}}\int_{\mathbb{R}^n} \mathcal{F}\left(\partial_{x_k}\varphi\right)(x)dx\\ =&\frac{1}{2\pi\mathbf{i}}\mathcal{F}^{-1}\left(\mathcal{F}\left(\partial_{x_k}\varphi\right)(x)\right)(0)\\ =&\frac{1}{2\pi\mathbf{i}}\partial_{x_k}\varphi(0).\end{aligned}$$

That is, in the sense of distributions,

$$\mathcal{F}\left(x^k\right) = \frac{1}{2\pi \mathbf{i}} \partial_{x_k} \delta_0.$$

We conclude: Fourier transform send polynomials to derivatives of dirac delta's and vice versa.

Theorem 17.25 (Fourier transform on $L^2(\mathbb{R}^n)$)**.** *Let $f \in L^2(\mathbb{R}^n)$.*

(1) $\mathcal{F}f \in L^2(\mathbb{R}^n)$. That is, the tempered distribution $\mathcal{F}(f)$ belongs to $L^2(\mathbb{R}^n)$, in the sense of Definition 17.18.
(2) $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a bounded linear operator. In particular, whenever $f_k \in C_c^\infty(\mathbb{R}^n)$ converges to $f \in L^2(\mathbb{R}^n)$ then there exists a subsequence $k_i \to \infty$

$$\mathcal{F}f(x) = \lim_{i\to\infty} \mathcal{F}f_{k_i}(x) \quad a.e.\ x \in \mathbb{R}^n.$$

(3) We have all the properties of Theorem 17.11 for $f, g \in L^2(\mathbb{R}^n)$. The Fourier inverse becomes

$$\mathcal{F}^{-1}\left(\mathcal{F}f)\right)(x) = \mathcal{F}\left(\mathcal{F}^{-1}f\right))(x) = f(x) \quad \textit{almost every } x \in \mathbb{R}^n$$

(4) $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is an isomorphic isometry, that is it satisfies

$$\|\mathcal{F}f\|_{L^2(\mathbb{R}^n)} = \|\mathcal{F}g\|_{L^2(\mathbb{R}^n)}.$$

Proof. (1) Observe that if $f \in L^2(\mathbb{R}^n)$ the tempered distribution satisfies by Hölder's inequality and Theorem 17.11.

$$\begin{aligned} |\mathcal{F}f[\varphi]| = |f[\mathcal{F}\varphi]| = & \left|\int_{\mathbb{R}^n} f\, \mathcal{F}\varphi\right| \\ \le & \|f\|_{L^2(\mathbb{R}^n)} \|\mathcal{F}\varphi\|_{L^2} \\ = & \|f\|_{L^2(\mathbb{R}^n)} \|\varphi\|_{L^2} \end{aligned}$$

By Proposition 17.21 we find that $\mathcal{F}f$ belongs to $L^2(\mathbb{R}^n)$ (observe that $2' = 2$). So we have

$$\mathcal{F}f[\varphi] = \int_{\mathbb{R}^n} \mathcal{F}f\, \varphi \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

Moreover,

$$\|\mathcal{F}f\|_{L^2(\mathbb{R}^n)} \le \|f\|_{L^2(\mathbb{R}^n)}. \tag{17.6}$$

Clearly $\mathcal{F}$ is linear, so (17.6) shows that $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a *bounded linear operator with norm ≤ 1.*

(2) Since $\mathcal{F} : L^2 \to L^2$ is a bounded linear operator, we have from Equation (17.6)

$$\|\mathcal{F}f - \mathcal{F}\tilde{f}\|_{L^2(\mathbb{R}^n)} \le \|f - g\|_{L^2(\mathbb{R}^n)}.$$

In particualr if $f_k \in C_c^\infty(\mathbb{R}^n)$ converges w.r.t L^2 to f

$$\|\mathcal{F}f - \mathcal{F}f_k\|_{L^2(\mathbb{R}^n)} \le \|f - f_k\|_{L^2(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

L^2-convergence implies a.e. convergence *up to subseqeuence*.

(3) Let $f, g \in L^2(\mathbb{R}^n)$ then there exist $f_k, g_k \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|f - f_k\|_{L^2(\mathbb{R}^n)}, \ \|g - g_k\|_{L^2(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

By the boundedness of $\mathcal{F}$ we then have that

$$\|\hat{f} - \hat{f}_k\|_{L^2(\mathbb{R}^n)}, \ \|\hat{g} - \hat{g}_k\|_{L^2(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

We now apply Theorem 17.11 to f_k and g_k (observe that $\hat{f}_k$, $\hat{g}_k \in \mathscr{S}(\mathbb{R}^n)$). Taking the limit it is then easy (by bilinearity) to see that Theorem 17.11 (1), (3), (4) hold for f and g. Again we conclude Theorem 17.11(5) from Theorem 17.11(3) by plugging in $f = g$.

For the Fourier inverse formula, up to taking a several subsequences we also have from (2) above and Theorem 17.11(2)

$$\mathcal{F}(\mathcal{F}^{-1}f)(x) = \lim_{i\to\infty} \mathcal{F}(\mathcal{F}^{-1}f_{k_i})(x) = \lim_{i\to\infty} f_{k_i}(x) = f(x) \quad \text{a.e. } x \in \mathbb{R}^n.$$

(4) is a direct consequence of Theorem 17.11 taking $f = g$.

□

Exercise 17.26. *Let $g \in C_c^\infty(\mathbb{R})$, and denote*

$$G(t) := \int_{-\infty}^t g(s)ds.$$

Show that G may not belong to $\mathscr{S}(\mathbb{R})$, but $G \in \mathscr{S}'(\mathbb{R})$. Show that

$$\mathcal{F}G(\tau) = \frac{g(\tau)}{2\pi \mathbf{i}\tau} + \pi g(0)\delta_{\tau=0}(\tau)$$

Hint: write $G(t)$ as a convolution of g and $\chi_{[0,\infty)}$.

Theorem 17.27. *Let $m \in L^\infty(\mathbb{R}^n)$ and define the following* multiplier operator *for $f \in L^2(\mathbb{R}^n)$.*

$$Tf(x) := \mathcal{F}^{-1}\left(m(\xi)\mathcal{F}f(\xi)\right).$$

Then f is a well defined linear operator and

$$\|Tf\|_{L^2(\mathbb{R}^n)} \le \|m\|_{L^\infty}\|f\|_{L^2(\mathbb{R}^n)}.$$

Remark 17.28.
- Actually one can show that a multiplier operator is bounded in $L^2(\mathbb{R}^n)$ if and only if $m \in L^\infty(\mathbb{R}^n)$.
- This gets way more complicated for $L^p(\mathbb{R}^n)$ (it is an ineed open question what are necessary *and* sufficient conditions): $m \in L^\infty$ is necessary so that T is bounded from L^p to L^p, but it does not suffice to ensure L^p-boundedness (essentially good decay at infinity and differentiability of m are sufficient). This is related to *Calderon-Zygmund theory* in Harmonic Analysis.
- Statements such as Theorem 17.27 are called *multiplier theorems* in Harmonic Analysis, famous ones are due to Marcinkiewicz, Mikhlin, Hörmander.

Proof. By Plancherel's theorem

$$\|Tf\|_{L^2(\mathbb{R}^n)} = \|\widehat{Tf}\|_{L^2(\mathbb{R}^n)} = \|m(\xi)\hat{f}\|_{L^2(\mathbb{R}^n)} \le \|m\|_{L^\infty}\|\hat{f}\|_{L^2(\mathbb{R}^n)} = \|m\|_{L^\infty}\|f\|_{L^2(\mathbb{R}^n)}.$$

□

Example 17.29. Consider $u \in L^p(\mathbb{R}^n)$ solves in distributional sense

$$\Delta u = f \quad \text{in } \mathbb{R}^n.$$

From our computations at the beginning of the

$$u = \mathcal{F}^{-1}\left(|\xi|^{-2}\mathcal{F}f\right)$$

So we have for any $\alpha, \beta \in \{1, \ldots, n\}$,

$$\partial_{\alpha\beta} u = c\mathcal{F}^{-1}\left(\frac{\xi^\alpha \xi^\beta}{|\xi|^2}\mathcal{F}f\right)$$

Clearly we have

$$\left|\frac{\xi^\alpha \xi^\beta}{|\xi|^2}\right| \lesssim 1.$$

So from Theorem 17.27 we conclude that actually

$$\|\partial_{\alpha\beta} u\|_{L^2(\mathbb{R}^n)} \lesssim \|f\|_{L^2(\mathbb{R}^n)}.$$

That is if $\Delta u = f$ we controll all the second derivatives of f (in L^2). In other words we have the (amazing) result

$$\|D^2 u\|_{L^2(\mathbb{R}^n)} \lesssim \|\mathrm{tr}(D^2 u)\|_{L^2(\mathbb{R}^n)} - -$$

and this even though $\mathrm{tr}(D^2u)$ only takes into account the diagonal of D^2u!!! □

If $\mathcal{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is bounded then surely $\mathcal{F} : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ is bounded? *WRONG!*

Theorem 17.30 (Fourier transform on L^p)**.** *Let $p \in [1, 2]$ and $p' = \frac{p}{p-1} \in [2, \infty]$. If $f \in L^p(\mathbb{R}^n)$ then $\hat{f} \in L^{p'}(\mathbb{R}^n)$ and*

$$\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)} \le \|f\|_{L^p(\mathbb{R}^n)}.$$

Proof. We will prove

$$\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)} \le C\|f\|_{L^p(\mathbb{R}^n)}$$

for a constant $C = C(p) > 0$. With complex interpolation one can show that $C = 1$.

For $\underline{p = 2,\ p' = 2}$ this is a consequence of Theorem 17.25.

For $\underline{p = 1,\ p' = \infty}$ it is a consequence of Fubini's theorem that distributional and pointwise Fourier transform coincide almost everywhere. Namely if $f \in L^1(\mathbb{R}^n)$ then

$$\int_{\mathbb{R}^n} \hat{f}\ \varphi = \int_{\mathbb{R}^n} f\ \hat{\varphi} \quad \forall \varphi \in C_c^\infty(\mathbb{R}^n).$$

So the claim follows from Lemma 17.8(1).

For some $p \in (1,2)$ the claim then follows from interpolation, Theorem 4.15. □

Remark 17.31. For $p > 2$ the result of Theorem 17.30 is false. Indeed, for any $p \in (2,\infty)$ there are functions $f \in L^p(\mathbb{R}^n)$ for which $\hat{f}$ (which is a tempered distribution) is *not* an even locally integrable function. See [Grafakos, 2014, Exercise 2.3.13]

We have observed above that $\mathcal{F}L^1 \subset L^\infty$. Actually we even have

Exercise 17.32. *Let $f \in L^1(\mathbb{R}^n)$ then $\hat{f}$ is continuous (in the sense that there is a representative of $\hat{f} \in L^\infty(\mathbb{R}^n)$ which is continuous).*

Hint: Write the integral of $\hat{f}(\xi) - \hat{f}(\eta)$ and use dominated convergence theorem.

Let us pretend we want to solve for a given $f \in L^p(\mathbb{R}^n)$ the *Laplace equation*

$$\Delta u(x) = f(x) \quad x \in \mathbb{R}^n,$$

where $\Delta = \sum_{i=1}^n \partial_i\partial_i$. Taking the Fourier transform we find that $\hat{u}$ should satisfy

$$c|\xi|^2\hat{u}(\xi) = \hat{f}(\xi),$$

that is u should be such that

$$\hat{u} := \frac{1}{c}|\xi|^{-2}\hat{f}(\xi)$$

That is we have the almos explicit representation

$$u = \mathcal{F}^{-1}\left(\frac{1}{c}|\xi|^{-2}\hat{f}(\xi)\right).$$

We can simplify this representation and obtain the *Newton potential* or *Riesz potential*

Theorem 17.33 (Fourier transform of $|x|^{\alpha-n}$)**.** *Let $f \in C_c^\infty(\mathbb{R}^n)$ and $\alpha \in (0,n)$. Then*

$$c_\alpha \mathcal{F}^{-1}\left(|\xi|^{-\alpha}\hat{f}(\xi)\right) = c_{n-\alpha}\int_{\mathbb{R}^n} |x-y|^{\alpha-n} f(y)dy.$$

Here $c_\alpha := \pi^{-\alpha/2}\Gamma(\alpha/2)$, where Γ denotes the Γ-function,

$$\Gamma(\alpha) = \int_0^\infty \mu^{\alpha-1}e^{-\mu}dx.$$

Proof. Since $f \in C_c^\infty(\mathbb{R}^n)$, $\hat{f}(\xi) \in \mathscr{S}(\mathbb{R}^n)$. The map $\xi \mapsto |\xi|^{-\alpha} \in L^1_{loc}(\mathbb{R}^n)$ if $\alpha < n$ (exercise! Polar coordinates), and moreover since $\alpha > 0$ the map $|\xi|^{-\alpha}\hat{f}(\xi)$ is a tempered distribution which is also integrable in $\mathbb{R}^n$. Thus $\mathcal{F}(|\xi|^{-\alpha}\hat{f}(\xi))$ is well-defined (as a distribution and a function).

We first claim that for any $\xi \neq 0$

$$c_\alpha|\xi|^{-\alpha} = \int_0^\infty e^{-\pi|\xi|^2\lambda}\lambda^{\frac{\alpha}{2}-1}d\lambda.$$

Indeed we have setting $\mu := \pi|\xi|^2\lambda$

$$\begin{aligned}
&\int_0^\infty e^{-\pi|\xi|^2\lambda}\lambda^{\frac{\alpha}{2}-1}d\lambda\\
=&\left(\pi|\xi|^2\right)^{1-\frac{\alpha}{2}}\int_0^\infty e^{-\pi|\xi|^2\lambda}\left(\pi|\xi|^2\lambda\right)^{\frac{\alpha}{2}-1}d\lambda\\
=&\left(\pi|\xi|^2\right)^{1-\frac{\alpha}{2}}\int_0^\infty e^{-\mu}\mu^{\frac{\alpha}{2}-1}\frac{d\mu}{\pi|\xi|^2}\\
=&\left(\pi|\xi|^2\right)^{-\frac{\alpha}{2}}\int_0^\infty e^{-\mu}\mu^{\frac{\alpha}{2}-1}d\mu\\
=&\left(\pi|\xi|^2\right)^{-\frac{\alpha}{2}}\Gamma(\alpha/2)\\
=&c_\alpha|\xi|^{-\alpha}.
\end{aligned}$$

Then we have by Fubini's theorem (applicable since we are integrable)

$$\begin{aligned}
c_\alpha\mathcal{F}^{-1}\left(|\xi|^{-\alpha}\hat{f}(\xi)\right)(x) &= \int_{\mathbb{R}^n}\int_0^\infty e^{-\pi|\xi|^2\lambda}\lambda^{\frac{\alpha}{2}-1}d\lambda\, e^{2\pi\mathbf{i}\langle\xi,x\rangle}\,\hat{f}(\xi)\,d\xi\\
&=\int_0^\infty \lambda^{\frac{\alpha}{2}-1}\left(\int_{\mathbb{R}^n} e^{2\pi\mathbf{i}\langle\xi,x\rangle}\,\left(e^{-\pi|\xi|^2\lambda}\hat{f}(\xi)\right)\right)d\xi d\lambda
\end{aligned}$$

Observe that the interior integral is the inverse Fourier transform of $e^{-\pi|\cdot|^2\lambda}f(\cdot)$, and the Fourier transform makes products into convolutions. That is

$$\begin{aligned}
&\left(\int_{\mathbb{R}^n} e^{2\pi\mathbf{i}\langle\xi,x\rangle}\,\left(e^{-\pi|\xi|^2\lambda}\hat{f}(\xi)\right)\right)d\xi\\
=&\mathcal{F}^{-1}\left(e^{-\pi|\cdot|^2\lambda}\hat{f}(\cdot)\right)(x)\\
=&\mathcal{F}^{-1}\left(e^{-\pi|\cdot|^2\lambda}\right)*\underbrace{\mathcal{F}^{-1}\hat{f}}_{=f}(x)\\
=&\mathcal{F}^{-1}\left(e^{-\pi|\cdot|^2\lambda}\right)*f(x)
\end{aligned}$$

Now with Example 17.5 and scaling properties of the Fourier transform

$$\mathcal{F}^{-1}\left(e^{-\pi|\cdot|^2\lambda}\right)(z)=\mathcal{F}^{-1}\left(e^{-\pi|\sqrt{\lambda}\cdot|^2}\right)(z)=\lambda^{-\frac{n}{2}}e^{-\pi\left|\frac{z}{\lambda}\right|^2},$$

so we arrive at

$$\begin{aligned}
c_\alpha\mathcal{F}^{-1}\left(|\xi|^{-\alpha}\hat{f}(\xi)\right)(x) &= \int_0^\infty \lambda^{\frac{\alpha}{2}-1}\int_{\mathbb{R}^n}\lambda^{-\frac{n}{2}}e^{-\pi\left|\frac{x-y}{\lambda}\right|^2}f(y)dy\,d\lambda\\
&=\int_{\mathbb{R}^n}\left(\int_0^\infty \lambda^{\frac{\alpha}{2}-1}\lambda^{-\frac{n}{2}}e^{-\pi\left|\frac{x-y}{\lambda}\right|^2}d\lambda\right)f(y)dy
\end{aligned}$$

As above we have

$$\left(\int_0^\infty \lambda^{\frac{\alpha}{2}-1}\lambda^{-\frac{n}{2}}e^{-\pi\left|\frac{x-y}{\lambda}\right|^2}d\lambda\right)=c_{n-\alpha}|x-y|^{\alpha-n},$$

and we can conclude. □

Remark 17.34.
- The formula of Theorem 17.33 holds also for $f \in L^p(\mathbb{R}^n)$, cf. [Lieb and Loss, 2001, Corollary 5.10]
- One can indeed compute that if $n \geq 3$ and $f \in L^p(\mathbb{R}^n)$, $p \in (1,\infty)$,
$$u(x) := \int_{\mathbb{R}^n} |x-y|^{2-n} f(y)dy.$$
then u is a distribution and there exists a constant C (explicitely computable) such that
$$\Delta u = Cf \quad \text{in } \mathcal{D}'(\mathbb{R}^n).$$

17.5. **Real Fourier transform.** The cis-formula says
$$e^{2\pi \mathbf{i}\langle \xi, x\rangle} = \cos(2\pi\langle \xi, x\rangle) + \mathbf{i}\sin(2\pi\langle \xi, x\rangle)$$
So we could have written (but won't)
$$\Re \mathcal{F} f(\xi) = \int_{\mathbb{R}^n} \cos(2\pi\langle \xi, x\rangle) f(x)\, dx$$
and
$$\mathrm{im} \mathcal{F} f(\xi) = \int_{\mathbb{R}^n} \sin(2\pi\langle \xi, x\rangle) f(x)\, dx$$

17.6. **Fourier transform for periodic functions.** The Fourier transform maps $L^2(\mathbb{R}^n)$ into $L^2(\mathbb{R}^n)$, and this makes us feel that $\mathcal{F}$ maps $L^2(\mathbb{R}^n)$ into *itself*. But this is not really the case, as the Heisenberg uncertainty principle illustrates: the domain $L^2(\mathbb{R}^n)$ ($\mathbb{R}^n$ is the geometric space) is different from the target $L^2(\mathbb{R}^n)$ (the phase space).

This becomes more visible if one considers Fourier transforms on periodic functions.

Definition 17.35 (periodic functions)**.** A function $f \in C^\infty_{per}([0,1]^n)$ is a function $f \in C^\infty(\mathbb{R}^n)$ with $f(x+k) = f(x)$ for any $k = (k_1,\ldots,k_n) \in \mathbb{Z}^n$.

The *torus* $\mathbb{T}^n := \mathbb{R}^n/\sim$ where $x \sim y$ iff $x = y+k$ for some $k \in \mathbb{Z}^n$. In other words, $\mathbb{T}^n = [0,1)^n$.

Any function $f \in L^p(\mathbb{T}^n)$ can be extended to a periodic function on $\mathbb{R}^n$ (remember we only need to define it a.e.)
$$\tilde{f}(x) := f(x-k) \quad \text{where } k \in \mathbb{Z}^n \text{ such that } x-k \in \mathbb{T}^n$$
Observe that $\tilde{f}$ probably is not in $L^p(\mathbb{R}^n)$ since it does not decay at infinity.

Exercise 17.36. *Let $p \in [1,\infty)$ and $f \in L^p([0,1])$. Show that for any $\varepsilon > 0$ there exists $f_\varepsilon \in C^0_{per}([0,1])$ such that $\|f - f_\varepsilon\|_{L^p(\mathbb{R}^n)} < \varepsilon$.*

Hint: The periodicity is the main issue, show that we can actually choose $f_\varepsilon \in C^\infty_c((0,1))$.

Definition 17.37. The *Fourier transform* for periodic function $f \in C^\infty_{per}([0,1]^n)$ is given by

$$\mathcal{F}f(k) := (\mathcal{L}^n([0,1]^n))^{-1} \int_{[0,1]^n} f(x) e^{-2\pi \mathbf{i} \langle x,k\rangle} dx \quad k \in \mathbb{Z}^n.$$

That is, $\mathcal{F}$ maps $f \in C^\infty_{per}([0,1]^n)$ into a sequence $(\mathcal{F}f(k))_{k\in\mathbb{Z}^n}$.

We will see: instead of mapping from $L^2(\mathbb{R}^n)$ to $L^2(\mathbb{R}^n)$, for periodic functions $\mathcal{F}$ maps from $L^2(\mathbb{R}^n)_{per}$ to $\ell^2(\mathbb{Z}^n)$.

Exercise 17.38.
- $\|\mathcal{F}f\|_{\ell^\infty(\mathbb{Z}^n)} \leq (\mathcal{L}^n([0,1]^n))^{-1} \|f\|_{L^1(\mathbb{T}^n)}$.
- $\mathcal{F}(\partial_{x^\alpha} f) = 2\pi \mathbf{i} k^\alpha \hat{f}(k)$
- *Show that for $\ell, k \in \mathbb{Z}^n$ we have*

$$(\mathcal{L}^n([0,1]^n))^{-1} \int_{[0,1]^n} e^{-2\pi \mathbf{i}\langle x,k\rangle} e^{2\pi\mathbf{i}\langle x,\ell\rangle} dx = \begin{cases} 1 & \ell = k \\ 0 & \ell \neq k. \end{cases}$$

Definition 17.39.
- A *trigonometric polynomial* P of degree $\leq N$ on $\mathbb{T}^n$ is a function of the form

$$P(x) := \sum_{k \in \{-N,\ldots,0,\ldots,N\}^n} c_k e^{2\pi\mathbf{i}\langle k,x\rangle}$$

  where $c_k \in \mathbb{C}$ are coefficients.
- The *N-th Fourier polynomial* of a map $f \in L^1(\mathbb{T}^n)$ is given by

$$S_N f(x) := \sum_{k \in \{-N,\ldots,0,\ldots,N\}^n} \mathcal{F}f(k) e^{2\pi\mathbf{i}\langle k,x\rangle}.$$

  Clearly $S_N f$ is a trigonometric polynomial.
- The (formal) *Fourier series* of a map $f \in L^1(\mathbb{T}^n)$ is given by

$$Sf(x) := \lim_{N\to\infty} S_N f(x),$$

  whenever that limit exists.
- More generally, the *inverse Fourier transform* for a sequence $(c_k)_{k\in\mathbb{Z}^n}$ is given by

$$\mathcal{F}^{-1}((c_k)_{k\in\mathbb{Z}})(x) := \sum_{k\in\mathbb{Z}^n} c_k e^{2\pi\mathbf{i}\langle x,k\rangle} := \lim_{N\to\infty} \sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} \mathcal{F}c_k e^{2\pi\mathbf{i}\langle k,x\rangle},$$

  again whenever this series converges.

Exercise 17.40. *Use the complex Stone-Weierstrass theorem (see Advanced Calculus), to show that for any $f \in C^0_{per}([0,1])$ and any $\varepsilon > 0$ there exists a trigonometric polynomial P_ε (of some finite order) such that*

$$\|f - P\|_{L^\infty([0,1])} < \varepsilon.$$

Definition 17.41 (The nuisance factor)**.** Define

$$\omega := \mathcal{L}^n(\mathbb{T}^n)^{-1}.$$

We want to show that $\mathcal{F}^{-1}\mathcal{F}f = f$. For this we first observe that $S_N[f]$ is the projection of f onto the space of trigonometric polynomials of degree $\leq N$.

Theorem 17.42. *Let $f \in L^2(\mathbb{T}^n)$ and $N \in \mathbb{N}$. Then we have for any trigonometric polynomial P of degree $\leq N$*

1. *$\omega\|f - S_N[f]\|^2_{L^2(\mathbb{T}^n)} \leq \omega\|f - P\|^2_{L^2(\mathbb{T}^n)}$, and equality holds if and only if $P = S_N f$.*
2. *$\omega\|f - S_N[f]\|^2_{L^2(\mathbb{T}^n)} = \omega\|f\|^2_{L^2(\mathbb{T}^n)} - \sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} |\hat{f}(k)|^2$.*
3. *In particular we have $\|S_N[f]\|_{L^2(\mathbb{T}^n)} \leq C\|f\|_{L^2(\mathbb{T}^n)}$ for some uniform constant C (depending on n, but not on N or f).*

Proof. For simplicity let us denote

$$e_k := e^{2\pi\mathbf{i}\langle k,x\rangle}$$

Denote by

$$\langle f, g\rangle := \omega\int_{\mathbb{T}^n} f(x)\overline{g}(x)dx$$

the scalar product. Then we have, Exercise 17.38,

$$\langle e_k, e_\ell\rangle = \delta_{k\ell}.$$

Let P be a generic trigonometric polynomial of degree at most N, i.e.

$$P(x) = \sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} c_k e_k.$$

Then we have

$$\begin{aligned}
\langle f-P, f-P\rangle =&\langle f,f\rangle + \langle P,P\rangle - \langle f,P\rangle - \overline{\langle f,P\rangle}\\
=&\omega\|f\|^2_{L^2} + \sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} |c_k|^2 - \sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} \left(\hat{f}(k)\overline{c_k} - \overline{\hat{f}(k)}c_k\right)\\
=&\omega\|f\|^2_{L^2} + \underbrace{\sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} |\hat{f}(k) - c_k|^2}_{\geq 0} - \left|\hat{f}(k)\right|^2
\end{aligned}$$

In particular we have (applying the above for $P = S_N[f]$),

$$\omega\|f - S_N f\|^2_{L^2} = \omega\|f\|^2_{L^2} - \sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} \left|\hat{f}(k)\right|^2.$$

Thus for any P,

$$\omega\|f-P\|^2_{L^2} = \omega\|f - S_N f\|^2_{L^2} + \underbrace{\sum_{k\in\{-N,\ldots,0,\ldots,N\}^n} |\hat{f}(k) - c_k|^2}_{\geq 0} \geq \|f - S_N f\|^2_{L^2}$$

and equality holds if and only if $P = S_N f$.

(3) follows from (1) via triangular inequality taking $P \equiv 0$. $\square$

Exercise 17.43 (Bessel). *Prove this corollary to Theorem 17.42. Let $f \in L^2(\mathbb{T}^n)$ then*

$$\sum_{k\in\mathbb{Z}^n} |\hat{f}(k)|^2 \le \omega \|f\|^2_{L^2(\mathbb{T}^n)}.$$

In particular the left-hand side series is absolutely convergent, i.e. $\mathcal{F}$ maps $L^2(\mathbb{T}^n)$ into $\ell^2(\mathbb{Z}^n)$.

One can then prove the following version of Parseval/Plancherell

Theorem 17.44 (Parseval). *If $f \in L^2(\mathbb{T}^n)$ then*

(1) $\lim_{N\to\infty} \|f - S_N[f]\|_{L^2} = 0$
(2) $\omega\|f\|^2_{L^2(\mathbb{T}^n)} = \|\mathcal{F}f\|^2_{\ell^2(\mathbb{Z}^n)}$
(3) More generally

$$\langle f, g\rangle_{L^2(\mathbb{T}^n)} = \langle \mathcal{F}f, \mathcal{F}g\rangle_{\ell^2(\mathbb{Z}^n)},$$

that is

$$\omega \int_{\mathbb{T}^n} f\overline{g} = \sum_{k\in\mathbb{Z}^n} \hat{f}(k)\,\overline{\hat{g}(k)}.$$

Proof. (1) First assume that $f \in C^0_{per}([0,1])$ and let $\varepsilon > 0$. By Exercise 17.40 there exists a trigonometric polynomial P of some degree N such that

$$\|f - P\|_{L^2(\mathbb{T}^n)} \le \mathcal{L}^n(\mathbb{T}^n)^{\frac{1}{2}} \|f - P\|_{L^\infty(\mathbb{T}^n)} \le \mathcal{L}^n(\mathbb{T}^n)^{\frac{1}{2}}\varepsilon.$$

By Theorem 17.42 we then have

$$\|f - S_N[f]\|_{L^2(\mathbb{T}^n)} \le \|f - P\|_{L^2(\mathbb{T}^n)} \le \mathcal{L}^n(\mathbb{T}^n)^{\frac{1}{2}}\varepsilon.$$

Moreover, for any $k \ge N$ we can apply again Theorem 17.42 to obtain

$$\|f - S_k[f]\|_{L^2(\mathbb{T}^n)} \le \|f - S_N[f]\|_{L^2(\mathbb{T}^n)} \le \mathcal{L}^n(\mathbb{T}^n)^{\frac{1}{2}}\varepsilon.$$

This implies the claim under the assumption that $f \in C^0_{per}([0,1])$.

If $f \in L^2([0,1])$ take f_ε from Exercise 17.36 such that

$$\|f - f_\varepsilon\|_{L^2([0,1])} < \varepsilon.$$

By Theorem 17.42(3) and linearity of S_N, we have for any $k \in \mathbb{N}$

$$\|f - f_\varepsilon\|_{L^2([0,1])} + \|S_k f - S_k f_\varepsilon\|_{L^2([0,1])} \le C\varepsilon.$$

So pick N for f_ε as above, then for any $k \ge N$ we have

$$\|f - S_k[f]\|_{L^2(\mathbb{T}^n)} \le \|f_\varepsilon - S_k[f_\varepsilon]\|_{L^2(\mathbb{T}^n)} + C\varepsilon \le \tilde{C}\varepsilon.$$

This proves (1).

(2) follows from (3)
(3) Exercise 17.45

□

Exercise 17.45. *Prove Theorem 17.44(3)*

17.7. Application: Basel Problem.

Exercise 17.46 (Solution of the Basel problem)**.** $\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}$

hint: Set

$$f(x) := \begin{cases} x & x \in [0, \frac{1}{2}] \\ x-1 & x \in (\frac{1}{2}, 1] \end{cases}$$

and use Parseval's theorem.

17.8. Application: Isoperimetric Problem - Hurwitz proof in 2D.

The *isoperimetric problem* also called *Dido's problem*[41] is the following question: given a fixed length, what is the shape of a set of maximal area whose boundary has said length. It's solution is the isoperimetric inequality: for any open set Ω with smooth boundary $\partial\Omega$,

$$n\mathcal{L}^n(\Omega)^{\frac{n}{n-1}}\mathcal{L}^n(B(0,1))^{\frac{1}{n}} \leq \mathcal{H}^{n-1}(\partial\Omega),$$

with equality if and only if Ω is a ball.

In two dimensions there is a beautiful proof due to Hurwitz (which sadly does not work in higher dimensions). See also Section 15.2

Two-dimensional proof. So let us consider a set $\Omega \subset \mathbb{R}^2 = \mathbb{C}$. Denote by $\gamma : [0,1] \to \mathbb{R}^2$, $\gamma(0) = \gamma(1)$, the boundary curve parametrizing $\partial\Omega$.

$$\mathcal{H}^1(\partial\Omega) = \text{Length}(\gamma) = \int_0^1 |\gamma'(t)| dt.$$

By rescaling we can assume w.l.o.g. that Length$(\gamma) = 1$. By a reparametrization we may also assume that $|\gamma'(x)| = 1$ for all x (this is called arclength parametrization).

Then we need to show

$$\mathcal{L}^2(\Omega) \leq \frac{1}{4\pi}.$$

We have by Plancherel's theorem, Theorem 17.44, and then Exercise 17.38,

$$\begin{aligned}\omega =& \omega \int_0^1 \underbrace{|\gamma'(t)|^2}_{=1} dt = \omega \int_0^1 |\gamma_1'(t)|^2 + |\gamma_2'(t)|^2 dt \\ =& \sum_{k\in\mathbb{Z}} |\widehat{\gamma_1'}(k)|^2 + |\widehat{\gamma_2'}(k)|^2 \\ =& 4\pi^2 \sum_{k\in\mathbb{Z}} |k|^2 \left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2\right)\end{aligned}$$

We record this for later

$$\sum_{k\in\mathbb{Z}} |k|^2 \left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2\right) = \frac{\omega}{4\pi^2} \tag{17.7}$$

[41]named after Queen Dido, founder of Carthage, who had to optimize the area of her city that she could surround with a Bull's Hide (which she had cut into strips)

On the other hand we have by the integration by parts formula

$$\mathcal{L}^2(\Omega) = \int_\Omega 1dx = \int_\Omega \partial_1 x^1 dx = \int_{\partial\Omega} \nu^1(t)\, x^1(t) d\mathcal{H}^1(t)$$

By the area formula ($\gamma : (0,1) \to \partial\Omega$ is a diffeomorphism) we have

$$\begin{aligned}&\int_{\partial\Omega} \nu^1(t) x^1(t) d\mathcal{H}^1(t)\\ =&\int_0^1 \nu^1(\gamma(t))\gamma^1(t) dt\end{aligned}$$

Now observe that ν is the unit normal, so it is orthogonal to $\gamma'(t)$ (which has norm one) so $\nu(\gamma(t)) = (\gamma_2'(t), -\gamma_1'(t))$ (assuming that γ is correctly parametrized, otherwise we use $\gamma(-t)$ to ensure the unit normal is outwards facing).

Thus,

$$\omega\mathcal{L}^2(\Omega) = \omega \int_0^1 \nu_1(t)\nu_2'(t)\, dt = \langle \nu_1, \nu_2' \rangle_{L^2(\mathbb{T})}.$$

Using again Plancherell we have

$$\omega\mathcal{L}^2(\Omega) = \sum_{k\in\mathbb{Z}} \widehat{\nu_1}(k)\overline{\widehat{\nu_2'}(k)}.$$

Since A is a real number, we can take the real part and have again with Exercise 17.38

$$\omega\mathcal{L}^2(\Omega) = \pi \sum_{k\in\mathbb{Z}} k2\Re\left(\widehat{\nu_1}(k)\overline{\widehat{\nu_2}(k)}\right)$$

Using this and (17.7) we have

$$\begin{aligned}&\omega\left(\frac{1}{4\pi^2} - \frac{\mathcal{L}^2(\Omega)}{\pi}\right)\\ =&\sum_{k\in\mathbb{Z}} |k|^2 \left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2\right) - k2\Re\left(\widehat{\nu_1}(k)\overline{\widehat{\nu_2}(k)}\right)\\ =&\sum_{k\in\mathbb{Z}} \left(|k|^2 - |k|\right)\left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2\right)\\ &+ \sum_{k\in\mathbb{Z}} |k| \left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2 - \operatorname{sgn}(k)2\Re\left(\widehat{\nu_1}(k)\overline{\widehat{\nu_2}(k)}\right)\right).\end{aligned}$$

Using that $|a \pm b|^2 = |a|^2 + |b|^2 \pm 2\Re(a\bar{b})$ we conclude that

$$\begin{aligned}&\omega\left(\frac{1}{4\pi^2} - \frac{\mathcal{L}^2(\Omega)}{\pi}\right)\\ =&\sum_{k\in\mathbb{Z}} \left(|k|^2 - |k|\right)\left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2\right)\\ &+ \sum_{k\in\mathbb{Z}} |k| \left|\widehat{\gamma_1}(k) - \operatorname{sgn}(k)\mathbf{i}\widehat{\gamma_2}(k)\right|^2\\ \geq&0,\end{aligned}$$

since $|k|^2 \geq |k|$ for all $k \in \mathbb{Z}$.

We readily conclude that

$$\left(\frac{1}{4\pi^2} - \frac{\mathcal{L}^2(\Omega)}{\pi}\right) \geq 0,$$

i.e.

$$\frac{1}{4\pi} \geq \mathcal{L}^2(\Omega).$$

Equality holds if and only if both

$$0 = \sum_{k\in\mathbb{Z}} \left(|k|^2 - |k|\right) \left(|\widehat{\gamma_1}(k)|^2 + |\widehat{\gamma_2}(k)|^2\right) \tag{17.8}$$

and

$$0 = \sum_{k\in\mathbb{Z}} |k| \left|\widehat{\gamma_1}(k) - \operatorname{sgn}(k)\mathbf{i}\widehat{\gamma_2}(k)\right|^2 \tag{17.9}$$

(17.8) implies (since $|k|^2 > |k|$ for $|k| > 1$), that

$$\hat{\gamma}_1(k) = \hat{\gamma}_2(k) = 0 \quad \forall k \in \mathbb{Z}\backslash\{0,-1,1\}.$$

Also, from (17.9) we conclude in particular

$$\widehat{\gamma_1}(1) = \mathbf{i}\widehat{\gamma_2}(1)$$

$$\widehat{\gamma_1}(-1) = -\mathbf{i}\widehat{\gamma_2}(-1)$$

which leads to

$$\widehat{\gamma_1}(\pm 1)^2 + \widehat{\gamma_2}(\pm 1)^2 = 0. \tag{17.10}$$

But since $\mathcal{F}^{-1}\mathcal{F}\gamma = \gamma$ we have

$$\gamma_i(t) = \widehat{\gamma}_i(-1)e^{-2\pi\mathbf{i}t} + \widehat{\gamma}_i(0) + \widehat{\gamma}_i(1)e^{2\pi\mathbf{i}t}, \quad i = -1, 0, 1.$$

Thus, for $p := (\widehat{\gamma_1}(0), \widehat{\gamma_2}(0))$ (observe the Fourier transform at 0 is a real number!) we have

$$\begin{aligned}
&|\gamma(t) - p|^2 \\
=& \left(\gamma_1(t) - \widehat{\gamma_1}(0)\right)^2 + \left(\gamma_2(t) - \widehat{\gamma_2}(0)\right)^2 \\
=& \sum_{i=1}^{2} \left(\widehat{\gamma}_i(-1)e^{-2\pi\mathbf{i}t} + \widehat{\gamma}_i(1)e^{2\pi\mathbf{i}t}\right)^2 \\
=& \left((\widehat{\gamma_1}(-1))^2 + (\widehat{\gamma_2}(-1))^2\right) e^{-4\pi\mathbf{i}t} + \left((\widehat{\gamma_1}(1))^2 + (\widehat{\gamma_2}(1))^2\right) e^{-4\pi\mathbf{i}t} \\
&+ 2\widehat{\gamma_1}(-1)\widehat{\gamma_1}(1) + 2\widehat{\gamma_2}(-1)\widehat{\gamma_2}(-1).
\end{aligned}$$

By (17.10) we conclude that

$$|\gamma(t) - p|^2 = +2\widehat{\gamma_1}(-1)\widehat{\gamma_1}(1) + 2\widehat{\gamma_2}(-1)\widehat{\gamma_2}(-1) \equiv \text{const}$$

That is γ parametrizes a circle. $\square$

17.9. **Discrete Fourier Transform and periodicity.** More generally one can define Fourier tranform on locally compact Abelian groups G, and $\mathcal{F}$ maps G into its *Pontryagin dual group* $\hat{G}$.

If $G = \mathbb{R}^n$ then $\hat{G} = \mathbb{R}^n$; if $G = \mathbb{T}^n$ then $\hat{G} = \mathbb{Z}^n$.

For example we can define the Fourier transform on $X = \{0, \dots, K-1\}$ where $K \in \mathbb{N}$.

Let $f : X \to \mathbb{C}$ then for $k \in X$ define (careful: the sign convention is not the same everywhere!)

$$\hat{f}(k) := \frac{1}{\sqrt{K}} \sum_{\ell=0}^{K-1} f(\ell) e^{-2\pi \mathbf{i} \frac{\ell\, k}{K}}$$

The discrete Fourier transform (there is an analogue version for the periodic Fourier transform on $L^2(\mathbb{T})$) is good in detecting periodicity. Namely assume that f is p-periodic, i.e. $f(x) = f(y)$ if $x \equiv y \mod p$. Here we assume that $\mu := \frac{K}{p} \in \mathbb{N}$.

Theorem 17.47. *Assume that f is p-periodic as above then $\hat{f}(k) \neq 0$ implies that k is an integer multiple of $\frac{K}{p}$.*

For the proof of Theorem 17.47 we need

Exercise 17.48. *Show the following*

(1) *Let $x \in \mathbb{R} \backslash \mathbb{Z}$, $N \in \mathbb{N}$. Then*
$$\sum_{\ell=0}^{N-1} e^{2\pi \mathbf{i} \ell x} = (\mathbf{i} \sin(\pi(N-1)x) + \cos(\pi(N-1)x)) \frac{\sin(N\pi x)}{\sin(\pi x)}$$
Hint: Write the sum as a geometric sum, $\sum_{\ell=0}^{N-1} \left(e^{2\pi \mathbf{i} x}\right)^k$.

(2) *In particular*
$$\sum_{\ell=0}^{N-1} e^{2\pi \mathbf{i} \ell x} = 0$$
whenever $x \in \mathbb{R} \backslash \mathbb{Z}$ but $Nx \in \mathbb{Z}$.

Proof of Theorem 17.47. Using periodicity we have

$$\begin{aligned}
\hat{f}(k) =& \frac{1}{\sqrt{K}} \sum_{\ell=0}^{K-1} f(\ell) e^{-2\pi \mathbf{i} \frac{\ell\, k}{K}} \\
=& \frac{1}{\sqrt{K}} \sum_{\ell=0}^{p-1} f(\ell) \sum_{j=0}^{\frac{K}{p}} e^{-2\pi \mathbf{i} \frac{(\ell + jp)\, k}{K}} \\
=& \frac{1}{\sqrt{K}} \sum_{\ell=0}^{p-1} f(\ell) e^{-2\pi \mathbf{i} \frac{\ell\, k}{K}} \sum_{j=0}^{\frac{K}{p}} e^{-2\pi \mathbf{i} \frac{(jp)\, k}{K}}
\end{aligned}$$

Set $L := \frac{K}{p}$ and let $k = \Lambda L + \mu$ for $\Lambda \in \mathbb{Z}$ and $0 \leq \mu < L$. Since $e^{-2\pi \mathbf{i} \Lambda j} = 1$ we have

$$\sum_{j=0}^{\frac{K}{p}} e^{-2\pi \mathbf{i} \frac{(jp)\,k}{K}} = \sum_{j=0}^{L} e^{-2\pi \mathbf{i} \frac{j\,k}{L}}$$

$$= \sum_{j=0}^{L} e^{-2\pi \mathbf{i} \frac{j\,\mu}{L}}$$

We see from Exercise 17.48 that whenever $\mu > 0$ the above sum is zero. That is, whenever k is not an integer multiple of $L = \frac{K}{p}$ then $\hat{f}(k) = 0$. □

Even with noise the Fourier transform is still good in guessing periodicity. This is used in cryptography, because the periodicity can be used to compute (or rather: guess) the prime factorization of an integer. *Shor's algorithm*, [Shor, 1997], is a version of this proposed for quantum computers, which supposedly can compute the (Quantum-)-Fourier transform extremely efficient, and thus break most of prime-number based cryptography systems.

17.10. **Further reading on applications.** [Stein and Shakarchi, 2003, Chapter 4]

17.11. **Sobolev spaces via Fourier transform.** Without going into details, the Fourier transform is also very useful to obtain a definition of Sobolev spaces on $\mathbb{R}^n$ (or a periodic Torus):

Recall that

$$\mathcal{F}(\partial_\alpha f)(\xi) = c\mathbf{i}\xi^\alpha \mathcal{F}(f)(\xi).$$

Thus,

$$\sum_{\alpha=1}^{n} \mathcal{F}(\partial_\alpha f)(\xi)\,\mathcal{F}(\partial_\alpha f)(\xi) = c\mathbf{i}|\xi|^2 |\mathcal{F}(f)(\xi)|^2.$$

So, by Plancherel/Parseval Theorem 17.11 we have

$$\|f\|_{L^2(\mathbb{R}^n)} + \|Df\|_{L^2(\mathbb{R}^n)} = \|\mathcal{F}f\|_{L^2(\mathbb{R}^n)} + C\||\xi|\mathcal{F}f(\xi)\|_{L^2(\mathbb{R}^n)}.$$

This holds at least if $f \in \mathscr{S}(\mathbb{R}^n)$. Indeed, using approximation by C_c^∞-functions we have the following (cf [Evans, 2010, p.297])

Theorem 17.49. *Let $f \in L^2(\mathbb{R}^n)$ then the following are equivalent for any $k \in \mathbb{N}$.*

(1) $f \in W^{k,2}(\mathbb{R}^n)$
(2) $(1 + |\xi|^k)\mathcal{F}f(\xi) \in L^2(\mathbb{R}^d)$.
(3) $(1 + |\xi|^2)^{\frac{k}{2}}\mathcal{F}f(\xi) \in L^2(\mathbb{R}^d)$.

In that case we have

$$\|f\|_{W^{k,2}(\mathbb{R}^n)} \approx \|(1 + |\xi|^2)^{\frac{k}{2}}\mathcal{F}f\|_{L^2(\mathbb{R}^n)} \approx \|(1 + |\xi|^k)\mathcal{F}f\|_{L^2(\mathbb{R}^n)}$$

We could be curious wether this holds for $L^p(\mathbb{R}^n)$, $p \in (1,\infty)$. But observe that $f \in L^p(\mathbb{R}^n)$ is not the same as $\mathcal{F}f \in L^{p'}(\mathbb{R}^n)$, we only have an embedding.

Instead let us have a look at the following reformulation of the above theorem.

Theorem 17.50. *Let $f \in L^2(\mathbb{R}^n)$ then the following are equivalent for any $k \in \mathbb{N}$.*

(1) $f \in W^{k,2}(\mathbb{R}^n)$
(2) $\mathcal{F}^{-1}\left((1+|\xi|^k)\mathcal{F}f\right) \in L^2(\mathbb{R}^d)$.
(3) $\mathcal{F}^{-1}\left((1+|\xi|^2)^{\frac{k}{2}}\mathcal{F}f\right) \in L^2(\mathbb{R}^d)$.

In that case we have

$$\|f\|_{W^{k,2}(\mathbb{R}^n)} \approx \|\mathcal{F}^{-1}\left((1+|\xi|^2)^{\frac{k}{2}}\mathcal{F}f\right)\|_{L^2(\mathbb{R}^n)} \approx \|\mathcal{F}^{-1}\left((1+|\xi|^k)\mathcal{F}f\right)\|_{L^2(\mathbb{R}^n)}$$

This version of the definition is indeed extendable to $p \in (1,\infty)$. Namely we have

Theorem 17.51. *Let $p \in (1,\infty)$. Let $f \in L^p(\mathbb{R}^n)$ then the following are equivalent for any $k \in \mathbb{N}$.*

(1) $f \in W^{k,p}(\mathbb{R}^n)$
(2) $\mathcal{F}^{-1}\left((1+|\xi|^k)\mathcal{F}f\right) \in L^p(\mathbb{R}^d)$.
(3) $\mathcal{F}^{-1}\left((1+|\xi|^2)^{\frac{k}{2}}\mathcal{F}f\right) \in L^p(\mathbb{R}^d)$.

In that case we have

$$\|f\|_{W^{k,p}(\mathbb{R}^n)} \approx \|\mathcal{F}^{-1}\left((1+|\xi|^2)^{\frac{k}{2}}\mathcal{F}f\right)\|_{L^p(\mathbb{R}^n)} \approx \|\mathcal{F}^{-1}\left((1+|\xi|^k)\mathcal{F}f\right)\|_{L^p(\mathbb{R}^n)}$$

If we can do that, then why not do this for $k \in [0,\infty)$ (instead of only $k \in \mathbb{N}$?)

Definition 17.52 (Bessel potential spaces)**.** Let $p \in (1,\infty)$, $s \in [0,\infty)$. We say $f \in L^p(\mathbb{R}^n)$ belongs to the *Bessel potential space* or (one of the many) *fractional Sobolev space*s $H^{s,p}(\mathbb{R}^n)$, iff

$$\mathcal{F}^{-1}\left((1+|\xi|^2)^{\frac{s}{2}}\mathcal{F}f\right) \in L^p(\mathbb{R}^d)$$

and we set

$$\|f\|_{H^{s,p}(\mathbb{R}^n)} := \|\mathcal{F}^{-1}\left((1+|\xi|^2)^{\frac{s}{2}}\mathcal{F}f\right)\|_{L^p(\mathbb{R}^n)}.$$

Equivalently we have

Theorem 17.53. *$f \in H^{s,p}(\mathbb{R}^n)$ if and only if $f \in L^p(\mathbb{R}^n)$ and the* fractional Laplacian

$$(-\Delta)^{\frac{s}{2}}f := \mathcal{F}^{-1}\left(|\xi|^s\mathcal{F}f\right) \in L^p(\mathbb{R}^n).$$

We then have

$$\|f\|_{H^{s,p}(\mathbb{R}^n)} \approx \|f\|_{L^p(\mathbb{R}^n)} + \|(-\Delta)^{\frac{f}{2}}\|_{L^p(\mathbb{R}^n)}.$$

In general the above theorem is not obvious at all, since we cannot take absolute values inside the Fourier transform. However for $p = 2$ we have Plancherel's theorem!

Exercise 17.54. *Show Theorem 17.53 for $p = 2$.*

The notion *fractional Laplacian* is justified by the following

Exercise 17.55. *Show that for $f \in \mathscr{S}(\mathbb{R}^n)$ and $s = 2$ we have*

$$(-\Delta)^{\frac{s}{2}} f = c(-\Delta) f,$$

where c is some constant and

$$-\Delta = \partial_{11} + \partial_{22} + \ldots + \partial_{nn}$$

is the Laplacian.

Observe that it is quite unclear how to define $H^{s,p}(\Omega)$ – and indeed there are different options, all with advantages or disadvantages. The most common one (if $\partial\Omega$ is smooth at least) is by extension.

Definition 17.56. We say that $f \in H^{s,p}(\Omega)$ iff there exist $F \in H^{s,p}(\mathbb{R}^n)$ such that $F\Big|_{\Omega} = f$. We sat

$$\|f\|_{H^{s,p}(\Omega)} := \inf_F \|F\|_{H^{s,p}(\mathbb{R}^n)}$$

where the infimum is taken over all F such that $F\Big|_{\Omega} = f$.

One might think another option would be to write $\|f\|_{L^p(\Omega)} + \|(-\Delta)^{\frac{s}{2}} f\|_{L^p(\Omega)}$ – but what is $(-\Delta)^{\frac{s}{2}} f$ if f is defined only in Ω?

For completeness let us mention another fractional Sobolev spcae, the Gagliardo-Slobodeckij (or sometimes: Besov)-space.

Definition 17.57. Let $s \in (0,1)$ and $p \in (1,\infty)$, $\Omega \subset \mathbb{R}^n$. Then $f \in W^{s,p}(\Omega)$ iff $f \in L^p(\Omega)$ and

$$[f]_{W^{s,p}(\Omega)} := \left(\int_\Omega \int_\Omega \frac{|f(x) - f(y)|^p}{|x-y|^{n+sp}} dx\, dy \right)^{\frac{1}{p}} < \infty.$$

We then set

$$\|f\|_{W^{s,p}(\Omega)} := \|f\|_{L^p(\Omega)} + [f]_{W^{s,p}(\Omega)}.$$

One can show that $W^{s,2}(\mathbb{R}^n) = H^{s,2}(\mathbb{R}^n)$, but in general $W^{s,p} \neq H^{s,p}$.

Can we define $f \in W^{s,p}(\mathbb{R}^n)$ also via the Fouuier transform? Yes we can but it gets crazy – it leads to the notion of Besov-spaces and Triebel-Lizorkin spaces.

Let $f \in \mathscr{S}(\mathbb{R}^n)$, and take any $\eta \in C_c^\infty(B(0,4))$, $\eta \equiv 1$ in $B(0,1)$ and set

$$p(x) := \eta(2x) - \eta(x),$$

and

$$p_j(x) := p(2^j x).$$

We then set the Littlewood-Paley projection

$$\Delta_j f := \mathcal{F}^{-1}(p_j(\xi)\mathcal{F}f).$$

Essentially Δ_j cuts off f at the frequency $|\xi| \approx 2^j$.

Then we define

$$[f]_{\dot{F}^{s,p}_q} := \| \left(\sum_{j\in\mathbb{Z}} 2^{jsq}|\Delta_j f|^q\right)^{\frac{1}{q}} \|_{L^p(\mathbb{R}^n)}$$

then f belongs to the Triebel space $F^{s,p}_q$ if $f \in L^p(\mathbb{R}^n)$ and $[f]_{\dot{F}^{s,p}_q} < \infty$. And we can also set

$$[f]_{\dot{B}^{s,p}_q} := \left(\sum_{j\in\mathbb{Z}} 2^{jsq}\|\Delta_j f\|^q_{L^p(\mathbb{R}^n)}\right)^{\frac{1}{q}}.$$

Then f belongs to the Besov space $B^{s,p}_q$ if $f \in L^p(\mathbb{R}^n)$ and $[f]_{B^{s,p}_q} < \infty$.

We then have

$$[f]_{F^{s,p}_p} = [f]_{B^{s,p}_p} \approx [f]_{W^{s,p}(\mathbb{R}^n)} \quad f \in \mathscr{S}(\mathbb{R}^n)$$

and (this is called Littlewood-Paley theorem):

$$[f]_{F^{s,p}_2} = \|(-\Delta)^{\frac{s}{2}} f\|_{L^p(\mathbb{R}^n)} \quad f \in \mathscr{S}(\mathbb{R}^n).$$

18. Topological Fixed Point Theorems

We first recall the *Banach Fixed Point theorem*

Theorem 18.1. *Let X be a complete metric space and suppose $T : X \to X$ is a contraction, i.e. there exists $\lambda \in [0,1)$ such that*

$$d(T(x), T(y)) \le \lambda d(x,y) \quad \forall x,y \in X.$$

Then T has a unique fixed point, i.e. there exists exactly one $\bar{x} \in X$ such that $T\bar{x} = \bar{x}$.

Moreover, for any $x \in X$,

$$\bar{x} = \lim_{n\to\infty} T^n(x),$$

where $T^n = \underbrace{T \circ T \circ \ldots \circ T}_{n \text{ times}} : X \to X$.

Banach Fixed Point theorem is very nice and useful, but the notion of contraction is a strong assumptions. In some situations where Banach Fixed Point theorems do not work (as we shall see below) topological fixed point theorems may become applicable.

Recall first the (finite-dimensional!) Brouwer Fixed point theorem, Theorem 14.2. We stated it on unit ball, but it can be extended to any compact convex set (which is nonempty).

Theorem 18.2 (Brouwer Fixed Point)**.** *Let $K \subset \mathbb{R}^n$ be a nonempty compact convex set. Then any $f \in C^0(K, K)$ has a fixed point, i.e. there exists $x \in K$ such that $f(x) = x$.*

Lemma 18.3. *Every compact convex set $K \subset \mathbb{R}^n$ that consists of at least two points is homeomorphic to a unit ball in $\mathbb{R}^\ell$ for some $\ell \in \{0, \dots, n\}$. That is there exists a bijective $\tau : K \to \overline{B_1(0)} \subset \mathbb{R}^\ell$ such that τ and τ^{-1} are both continuous.*

Proof. Assume first that K has nonempty interior, i.e. there exists some $x_0 \in K$ and $r > 0$ such that $B(x_0, r) > 0$ then we can w.l.o.g. assume $x_0 = 0$ and set ρ to be the *Minkowski functional*, see Proof of Theorem 10.21, i.e.

$$\rho(x) := \inf \left\{ t > 0 : \quad \frac{1}{t} x \in K \right\}.$$

Then $\rho(x) < \infty$ since for each $x \in \mathbb{R}^n$ we have $\frac{1}{2} r \frac{x}{|x|} \in B(0, r) \subset K$. Since K is bounded, we have $\rho(x) > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$ and $\rho(0) = 0$.

On the other hand (cf. again proof of Theorem 10.21) we have

$$K = \{x \in \mathbb{R}^n : \rho(x) \le 1\},$$

where we use convexity and compactness of K. We also have (yet again see Theorem 10.21) $\rho(\lambda x) = \lambda \rho(x)$ and $\rho(x + y) \le \rho(x) + \rho(y)$.

We then set

$$\tau(x) := \begin{cases} x \frac{\rho(x)}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

We then have

$$K = \{x : \rho(x) \le 1\} = \{x : |\frac{x}{|x|} \rho(x)| \le 1\} = \{x : |\tau(x)| \le 1\}$$

That is $\rho : K \to \overline{B(0,1)}$ is surjective. Also, $\tau : K \to B(0, 1)$. ρ is injective: indeed assume

$$\tau(x) = \tau(y)$$

Clearly $\tau(x) = \tau(y) = 0$ if and only if $x = y = 0$ so we may assume $x, y \neq 0$. Then $|\tau(x)| = |\tau(y)|$ so $\rho(x) = \rho(y)$. Thus $\tau(x) = \tau(y)$ implies $\frac{x}{|x|} = \frac{y}{|y|}$ that is $x = \lambda y$ where $\lambda > 0$. But then $\rho(x) = \lambda \rho(y)$ which implies $\lambda = 1$ and thus $x = y$.

So $\tau : K \to \overline{B(0,1)}$ is bijective.

As for continuity let $x \in K$. First assume $x = 0$ then $\tau(x) = 0$. Recall that we know that $B(0, r) \subset K$ so if $0 < |y| < \frac{r}{2}$ we have $|\tau(y)| = 2r|y|\rho(y/(2r|y|)) \le 2r|y|$. In particular $|\tau(y)| \ll 1$ if $|y| \ll 1$, so τ is continuous at 0.

Next fix $x \neq 0$. Then $y \mapsto \frac{y}{|y|}$ is a Lipschitz function for all $y \approx x$ (since $y \neq 0$) and $\rho(x)$ is Lipschitz by reverse triangular inequality. So $y \mapsto \tau(y)$ is Lipschitz as a product of two Lipschitz functions for $y \approx x$.

We have by now shown that $\tau : K \to \overline{B(0,1)}$ is a continuous bijection, it remains to show that $\tau^{-1} : \overline{B(0,1)} \to K$ is continuous. However observe that

$$\tau^{-1}(x) := \begin{cases} x\frac{|x|}{\rho(x)} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

For the same reason as before, τ^{-1} is continuous: around any $x \neq 0$ it is a Lipschitz function since $\rho(x) \neq 0$. And for $x = 0$ we see that for $\tau^{-1}(y) = y\rho(\frac{y}{|y|})$ which since as a continuous function ρ is bounded on the compact set $\partial B(0,1)$ we have

$$|\tau^{-1}(y)| \leq |y| \sup_{|z|=1} \rho(z) \ll 1 \quad \text{if } |y| \ll 1.$$

This settles the claim if we assume $\underline{K \text{ has nonempty interior}}$.

However in general K may have no nonempty interior. But then it must be lower dimensional. Without loss of generality we again may assume that $0 \in K$.

Let

$$X := span\{K\} = \left\{ x \in \mathbb{R}^n : x = \sum_{i=1}^n \lambda_i x_i \quad \text{for some } (x_i)_{i=1}^n \in K \text{ and } (\lambda_i)_{i=1}^n \subset \mathbb{R} \right\}.$$

Clearly X is a linear subspace of $\mathbb{R}^n$ and $K \subset X$. If $\dim X = n$ then there must be n linear independent $x_i \in K \setminus \{0\}$, $i = 1, \ldots, n$. But then their convex hull

$$\{x_1, \ldots, x_n\}^{conv} := \{\sum_{i=1}^n \lambda_i x_i \quad \text{for any } \lambda_i \in [0,1] \text{ with } \sum_{i=1}^n \lambda_i = 1\} \subset K$$

contains an open neighborhood of the origin, i.e. K has nonempty interior in $\mathbb{R}^n$.

Thus, if the dimension of X must be strictly less than n (or K has nonempty interior). Let $\dim X = \ell$, and consider $h : \mathbb{R}^n \to \mathbb{R}^\ell$ the projection that maps X into $\mathbb{R}^\ell$ – this is a smooth diffeomorphism as a map $h : X \to \mathbb{R}^\ell$. In particular $h : K \to h(K)$ is smooth, and $h(K)$ is still convex. If $\ell \geq 1$ then we are sure that $h(K)$ has nonempty interior in $\mathbb{R}^\ell$ and we can conclude.

The only case left is if $\ell = 0$, then K consists of only one point which is excluded by assumption.

□

Exercise 18.4. *Prove Theorem 18.2 by combining Lemma 18.3 and the Brouwer Fixed Point theorem on a ball, Theorem 14.2.*

The *Brouwer Fixed Point Theorem* Theorem 18.2 can be extended to infinite dimensional spaces. One of these extension is the *Schauder Fixed Point Theorem*.

Theorem 18.5. *Let X be a Banach space and $K \subset X$ be compact, convex and nonempty. Let $F : K \to K$ be any continuous mapping (F does not need to be linear!). Then F has a fixed point, i.e. there exists $x \in K$ with $F(x) = x$.*

Proof. Fix $n \in \mathbb{N}$ and cover K by finitely many balls $B(x_k, \frac{1}{n})$, $k = 1, \ldots, K(n)$, and $x_k \in K$. Consider the convex hull of the centers of these balls

$$K^n := \{x_1, \ldots, x_{K(n)}\}^{conv} := \{ \sum_{i=1}^{K(n)} \lambda_i x_i \quad \text{for any } \lambda_i \in [0,1] \text{ with } \sum_{i=1}^{n} \lambda_i = 1\} \subset K.$$

K^n is a convex, finite dimensional set – and compact (since it is bounded and closed) – so we could apply Brouwer Fixed Point Theorem, Theorem 18.2. For this, we need to somehow restrict F to K^n but make sure that this restriction maps K^n into K^n.

For $x \in K$ set

$$G_n(x) := \frac{1}{\sum_{i=1}^{K(n)} \operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right)} \sum_{i=1}^{K(n)} \operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right) x^i.$$

Observe that for any $x \in K$ at least one of the summands must be nonzero, and the distance function is Lipschitz continuous – so $G_n(x)$ is a continuous function on K.

Also, if we set for some $x \in K$

$$\lambda_i := \frac{1}{\sum_{j=1}^{K(n)} \operatorname{dist}\left(x, K \setminus B(x_j, \frac{1}{n})\right)} \operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right)$$

then $\sum_i \lambda_i = 1$, $\lambda_i \in (0,1)$, and thus

$$G_n(x) \in K^n \quad \forall x \in K.$$

Moreover we have G_n is not too far away from the identity, indeed for any $x \in K$,

$$\begin{aligned}
\|G_n(x) - x\| \leq & \frac{1}{\sum_{i=1}^{K(n)} \operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right)} \sum_{i=1}^{K(n)} \underbrace{\operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right)}_{=0 \text{ unless } \|x^i - x\| \leq \frac{1}{n}} \|x^i - x\| \\
\leq & \frac{\sum_{i=1}^{K(n)} \operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right)}{\sum_{i=1}^{K(n)} \operatorname{dist}\left(x, K \setminus B(x_i, \frac{1}{n})\right)} \frac{1}{n} = \frac{1}{n}.
\end{aligned} \tag{18.1}$$

So if we set

$$F_n(x) := G_n(F(x))$$

Then $F_n : K \to K$ is continuous, and since G_n maps K into K^n we have

$$F_k(x) : K^n \to K^n$$

is continuous. By Brouwer's Fixed point theorem (since K^n is finite dimensional), Theorem 18.2, we conclude that F_n has at least one fixed point $x_n \in K^n$, i.e. $F_n(x_n) = x_n$.

In particular $(x_n)_{n\in\mathbb{N}} \in K$ and up to passing to a subsequence we may assume that x_n converges to some $x \in K$. We then have by continuity of F,

$$\begin{aligned}\|x - F(x)\| &= \lim_{n\to\infty} \|x_n - F(x_n)\| = \lim_{n\to\infty} \|F_n(x_n) - F(x_n)\| \\ &= \lim_{n\to\infty} \|G_n(F(x_n)) - F(x_n)\| \\ &\overset{(18.1)}{\leq} \lim_{n\to\infty} \frac{1}{n} = 0.\end{aligned}$$

Thus $F(x) = x$ and we can conclude. □

We want to slightly weaken the conditions of Theorem 18.5, namely we are going to show

Corollary 18.6. *Let X be a Banach space and $K \subset X$ be closed, convex and nonempty set. Let $F : K \to K$ be any continuous mapping and assume that $F(K)$ is precompact, i.e. $\overline{F(K)}$ is compact. Then F has a fixed point, i.e. there exists $x \in K$ with $F(x) = x$.*

We need the following result about convex sets.

Lemma 18.7. *Let X be a Banach space and $A \subset X$ be precompact, i.e. assume $\overline{A}$ is compact. Then the closure of the convex hull of A, $\overline{A^{conv}}$, is compact and convex.*

Proof. It is easy to show that $\overline{A^{conv}}$ is convex, we focus on compactness.

Fix any $\varepsilon > 0$, then we can cover A by finitely many balls $B(x_i, \varepsilon)$, $i = 1, \ldots, N(\varepsilon)$. Then

$$A^{conv} \subset \{x_1, \ldots, x_{N(\varepsilon)}\}^{conv} + B(0, \varepsilon).$$

On the other hand, $\{x_1, \ldots, x_{N(\varepsilon)}\}^{conv}$ is compact (as a finite dimensional closed and bounded set). So we can cover it by finitely many balls $B(x_i, \varepsilon)$, $i = N(\varepsilon)+1, \ldots, M(\varepsilon, x_1, \ldots, x_N(\varepsilon))$. That is, for any $\varepsilon > 0$ there exists $M = M(\varepsilon)$ such that

$$A^{conv} \subset \bigcup_{i=1}^{M} B(x_i, \varepsilon).$$

In particular, for any $\varepsilon > 0$ there exists $M = M(\varepsilon)$ such that

$$\overline{A^{conv}} \subset \bigcup_{i=1}^{M} B(x_i, 2\varepsilon).$$

This property is called *totally bounded* and together with the closedness of $B := \overline{A^{conv}}$ and completeness of the space it implies compactness: take any sequence $(z_i)_{i\in\mathbb{N}} \subset B = \overline{A^{conv}}$. Take $\varepsilon := \frac{1}{2}$ from the finite covering above there must be some $\bar{x}$ such that infinitely many sequence elements lie in $B(\bar{x}, 1)$. We keep the first sequence element of $(z_i)_i$ and then disgard all the sequence elements not in $B(\bar{x}_1, 1)$. Taking $\varepsilon = \frac{1}{4}$ we repeat the argument: infinitely many subsequence elements must belong to some $B(\bar{x}_2, \frac{1}{2}) \cap K$, we keep the first and second subsequence element, and disgard then all elements not in $B(\bar{x}_2, \frac{1}{2})$. By this we obtain a subsqeuence of which for each $k \in \mathbb{N}$ all but finitely many sequence elements lie in some $B(\bar{x}_k, \frac{1}{k}) \cap B$ – implying that the sequence is Cauchy, hence by completeness of X convergent, hence by closedness of B has a limit $z \in B$. Thus B is compact. □

Proof of Corollary 18.6. Consider

$$\tilde{K} := \overline{F(K)^{conv}}$$

Since K is convex, $F(K)^{conv} \subset K$. Since K is closed we have that $\tilde{K} \subset K$. By Lemma 18.7, $\tilde{K}$ is compact. Moreover we can restrict F to $\tilde{K}$ and obtain a continuous mapping $F\Big|_{\tilde{K}}$: $\tilde{K} \to \tilde{K}$ – By Theorem 18.5 F has a fixed point in $\tilde{K} \subset K$. □

A corollary of Corollary 18.6 is the following

Corollary 18.8. *Let X be a Banach space and assume let $B = B(0,1)$ be the open unit ball of X. Let $G : \overline{B} \to X$ with*

- *$G(\partial B) \subset B$, and*
- *$\overline{G(\overline{B})}$ is compact.*

Then G has a fixed point in $B(0,1)$, i.e. there exists x with $\|x\|<1$ such that $G(x) = x$.

Proof. We want to apply Corollary 18.6, however G maps into X not $\overline{B}$. This can be mitigated by a projection into $\overline{B}$:

Set

$$F(x) := \begin{cases} G(x) & \text{if } \|G(x)\| \leq 1 \\ \frac{G(x)}{\|G(x)\|} & \text{if } \|G(x)\| > 1. \end{cases}$$

Then $F : \overline{B} \to \overline{B}$ is a continuous map. Since $\overline{G(\overline{B})}$ is compact, so is $\overline{F(\overline{B})}$ (continuous map maps compact sets into compact sets, and the projection into $\overline{B}$ is continuous!).

So F has a fixed point, i.e. there exists $x \in \overline{B}$ with $F(x) = x$.

Assume that $\|x\| = 1$. Since by assumption $G(\partial B) \subset B$ we have $\|G(x)\| < 1$ thus $x = F(x) = G(x)$ which implies $\|x\| < 1$ – contradiction. So we know that $\|x\| < 1$.

But then $\|F(x)\| = \|x\| < 1$ and thus $G(x) = F(x) = x$. □

From the above we can deduce the the *Schaefer's fixed point theorem* also known as the *Leray-Schauder theorem*. Its main advantage is that we do not have to identify a convex domain for F.

Theorem 18.9 (Leray-Schauder)**.** *Let X be a Banach space and assume that $F : X \times [0,1] \to X$ be a continuous and compact mapping, in the sense that for any bounded set $A \subset [0,1] \times X$ we have $\overline{F(A)} \subset X$ is compact (i.e. $F(A)$ is precompact).*

- *$F(x,0) \equiv 0$ for all $x \in X$, and*
- *there is a constant $M > 0$ such that for all*
$$x \in X \text{ such that there exists } \mu \in [0,1] \text{ with } x = F(x,\mu)$$
we have $\|x\| \leq M$.

Then $F(\cdot,1)$ has a fixed point, i.e. there exists $x \in X$ such that $F(x,1) = x$.

Leray-Schauder is a homotopy argument from the constant map (every point is a fixed point).

Proof. Without loss of generality, $M = \frac{1}{2}$, otherwise we apply the argument below to
$$\tilde{F}(x,\mu) := \frac{1}{2M} F(2Mx,\mu).$$
Indeed, if $x \in X$ is a fixed point of $\tilde{F}$ for some $\mu \in [0,1]$, i.e. if $\tilde{F}(x,\mu) = x$, then
$$2Mx = F(2Mx,\mu)$$
and thus by assumption
$$2M\|x\| \leq M, \quad \Leftrightarrow \|x\| < \frac{1}{2}.$$
And $\tilde{F}$ still satisfies all the previous assumptions.

That is, from now we may assume that

$$\text{if } \mu \in [0,1] \text{ and } x = F(x,\mu) \quad \text{then} \|x\| < \frac{1}{2}. \tag{18.2}$$

For $\varepsilon \in (0,1]$ we define the map $G_\varepsilon : \overline{B(0,1)} \to X$ by
$$G_\varepsilon(x) := \begin{cases} F(\frac{x}{(1-\varepsilon)}, 1) & \text{if } \|x\| < 1-\varepsilon \\ F\left(\frac{x}{\|x\|}, \frac{1-\|x\|}{\varepsilon}\right) & \text{if } 1-\varepsilon \leq \|x\| \leq 1 \end{cases}$$
Then $G_\varepsilon : \overline{B(0,1)} \to X$ is continuous. Since $\overline{B(0,1)}$ is bounded and F is a compact map, we also have that $\overline{F(\overline{B(0,1)})}$ is compact. Lastly, if $\|x\| = 1$ then $G_\varepsilon(x) = F(x,0) = 0$. That is $G_\varepsilon(\partial B(0,1)) = \{0\} \subset B(0,1)$[42].

We can apply Corollary 18.8 and find that for each $\varepsilon \in (0,1]$ there must be some x_ε with $\|x_\varepsilon\| < 1$ with
$$G_\varepsilon(x_\varepsilon) = x_\varepsilon.$$
Apply this result to $\varepsilon = \frac{1}{k}$, then we find x_k, $\|x_k\| < 1$, with
$$G_{\frac{1}{k}}(x_k) = x_k \quad k \in \mathbb{N}$$

[42]Here we could have relaxed the assumption $F(x,0) \equiv 0$ a bit in the statement of the theorem

We set

$$\mu_k := \begin{cases} 1 & \text{if } \|x_k\| < 1 - \frac{1}{k} \\ k(1 - \|x_k\|) & \text{if } 1 - \frac{1}{k} \le \|x_k\| \le 1 \end{cases}$$

We then have

$$x_k = \begin{cases} F(\frac{x_k}{(1-\frac{1}{k})}, \mu_k) & \text{if } \|x_k\| < 1 - \frac{1}{k} \\ F\left(\frac{x_k}{\|x_k\|}, \mu_k\right) & \text{if } 1 - \frac{1}{k} \le \|x_k\| \le 1 \end{cases} \tag{18.3}$$

In particular $x_k \subset F(\overline{B(0,1)}, [0,1])$, so since F is a compact operator, we may choose a subsequence $k \to \infty$ (not relabeled) such that $x_k \xrightarrow{k\to\infty} \bar{x}$ in $\overline{B(0,1)} \subset X$. We may choose a further subsequence (again not relabeled) so that $\mu_k \xrightarrow{k\to\infty} \mu \in [0,1]$.

We claim that $\mu = 1$. Indeed, if this is not the case and $\mu < 1$ this must mean that for all but finitely many k we have $\mu_k < 1$ and thus $\|x_k\| > 1 - \frac{1}{k}$, so when passing to the limit we have $\|\bar{x}\| = 1$ and

$$\bar{x} = \lim_{k\to\infty} x_k = \lim_{k\to\infty} F\left(\frac{x_k}{\|x_k\|}, \mu_k\right) = F(\bar{x}, \mu)$$

But this is a contradiction to (18.2). So indeed $\mu = 1$.

Next we pass to the limit in the (18.3) and we have

$$\bar{x} = F(\bar{x}, 1) \quad \text{ or } \bar{x} = F(\frac{\bar{x}}{\|x\|}, 1)$$

If $\|\bar{x}\| = 1$ the two options coincide, however this cannot happen do to (18.2).

So we have $\|\bar{x}\| < 1$. Then we have again for all but finitely many k that $\|x_k\| < 1 - \frac{1}{k}$, so

$$x_k = F(\frac{x_k}{(1-\frac{1}{k})}, \mu_k)$$

and passing to the limit we have

$$\bar{x} = F(\bar{x}, 1)$$

That is $\bar{x}$ is the fixed point we were looking for.

We can conclude. □

Leray-Schauder is also often written in the following way:

Corollary 18.10 (Leray-Schauder). *Let $F : X \to X$ be a continuous and compact mapping of a Banach space X into itself, such that the set*

$$\{x \in X : \quad x = \mu F x \text{ for some } 0 < \mu \le 1\}$$

is bounded, i.e. there exists $M > 0$ such that whenever for some $\mu \in [0,1]$ and some $x \in X$ we have $x = \mu F x$ then

$$\|x\|_X \le M.$$

Then F has a fixed point.

Exercise 18.11. *Show that Corollary 18.10 is indeed a consequence of Theorem 18.9.*

The main point of Corollary 18.10 is that we can assume we have fixed points for μF, show their boundedness (a priori estimates), if we can prove appropriate estimates for solutions of a non- linear PDE, under the assumption that such solutions exist, then in fact these solutions do exist". This is the method of a priori* estimates. [] The advantage of *Schaefer's theorem* over Schauder's for applications is that we do not have to identify an explicit convex, compact set.

Observe that we get no uniqueness from Leray-Schauder.

18.1. **Fixed point theorems applied in PDE.** In PDE, fixed point theorems are often used to show existence of solutions which are distortions of "easy" PDE. Banach Fixed Point can deal with *small* perturbations, Leray Schauder is good for *compact distortions*.

Example 18.12 (Application of the Banach Fixed Point)**.** Let B be the unit ball in $\mathbb{R}^n$. Let $F : \mathbb{R}^n \to \mathbb{R}$ be a uniformly Lipschitz map, i.e. assume there exists some $\Lambda > 0$ such that

$$|F(A) - F(B)| \leq \Lambda |A - B| \quad \forall A, B \in \mathbb{R}^n.$$

Then there exists an $\varepsilon_0 > 0$ (depending on n and Λ) such that for each $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ there exists a unique solution $u \in H^1_0(B)$ to

$$\begin{cases} \Delta u = g - \varepsilon F(Du) & \text{in } B. \\ u = 0 & \text{on } \partial B \end{cases}$$

So the point is that $\Delta u = g$ we can easily solve, and $\Delta u + \varepsilon F(Du)$ is "not too far". Observe that since F could be nonlinear, it is difficult to obtain a solution variationally or by linear approximations (F need only be Lipschitz!). Here Banach Fixed Point, Theorem 18.1, comes into play. The point why our argument works is because $F(Du)$ is *lower order* (Δ has differential order two, Du is order one).

We need to construct T which is so that if it has a fixed point u than u is the solution we want.

We define $T : H^1_0(\Omega) \to H^1_0(\Omega)$ via

$$\Delta T u = g - \varepsilon F(Du).$$

I.e. given $u \in H^1_0(B)$ solve the equation $\Delta w = g - \varepsilon F(Du)$ (observe the right-hand side is in L^2 so we can easily find a unique solution $w \in H^1_0(\Omega)$) and call $Tu := w$. If it happens that we find a fixed point u, i.e. $Tu = u$, then u is indeed the solution we wanted.

Observe that $T : H^1_0(\Omega) \to H^1_0(\Omega)$ is *nonlinear* if F is nonlinear.

We need to show that Tu is a contraction: For this we consider the equation for $Tu - Tv$.

$$\Delta(Tu - Tv) = \varepsilon \left(F(Dv) - F(Du) \right)$$

We need some estimates, the trick is often to multiply this equation with $Tu - Tv$ and integrate by parts, i.e. (here we use that $Tu - Tv = 0$ on $\partial\Omega$)

$$\int_B \Delta(Tu - Tv)(Tu - Tv) = -\int D(Tu - Tv) \cdot D(Tu - Tv) = -\|D(Tu - Tv)\|^2_{L^2(B)}.$$

On the other hand we have the PDE, so we get

$$-\|D(Tu - Tv)\|^2_{L^2(B)} = \varepsilon \int_B (F(Dv) - F(Du))(Tu - Tv).$$

Thus, by Hölder's inequality,

$$\|D(Tu - Tv)\|^2_{L^2(B)} \leq |\varepsilon| \|F(Du) - F(Dv)\|_{L^2(B)} \|Tu - Tv\|_{L^2(B)}.$$

By Poincaré inequality, Corollary 13.38, (again using the boundary being zero) we then have

$$\|Tu - Tv\|^2_{H^1(B)} \leq C|\varepsilon| \|F(Du) - F(Dv)\|_{L^2(B)} \|Tu - Tv\|_{H^1(B)}.$$

Dividing both sided by $\|Tu - Tv\|_{H^1(B)}$ (if it is zero there is nothing to show) we obtain that (here we use the Lipschitz assumption on F)

$$\|Tu - Tv\|_{H^1(B)} \leq C|\varepsilon| \|F(Du) - F(Dv)\|_{L^2(B)} \leq C|\varepsilon|\Lambda \|Du - Dv\|_{L^2(B)} \leq C|\varepsilon|\Lambda \|u - v\|_{H^1(B)}.$$

Clearly if $|\varepsilon| < \varepsilon_0$ and ε_0 is so small such that $C\Lambda|\varepsilon| < \frac{1}{2}$ we get

$$\|Tu - Tv\|_{H^1(B)} \leq \frac{1}{2} \|u - v\|_{H^1(B)} \quad \forall u, v \in H^1_0(B),$$

thus T is the required contraction and thus it has a unique fixed point.

Applying *Banach Fixed Point theorem*, Theorem 18.1, is usually preferable to Leray-Schauder, since it provides automatically uniqueness, and a method of construction (the fixed point is given by $u = \lim_{n \to \infty} T^n u$). Leray-Schauder doesn't provide any hint where the fixed point is, but it works e.g. for non-Lipschitz maps.

Example 18.13 (Leray-Schauder)**.** Leray-Schauder, Corollary 18.10, is very helpful to solve compact distortions of nice PDE, for which we nevertheless cannot use Banach Fixed Point theory Theorem 18.1 (which gives uniqueness!).

As an example let us try to find a solution $u \in H^1_0(B)$ (B is the unit ball in $\mathbb{R}^d$) to

$$\Delta u = g + |Du|^\theta \quad \text{in } B,$$

where $g \in L^2(B)$ is a given fixed function and $\theta \in (0, 1)$.

Again we define the operator Tu via

$$\Delta(Tu) = g + |Du|^\theta.$$

If $u \in H^1(B)$ then $|Du| \in L^2(B)$ so $|Du|^\theta \in L^{\frac{2}{\theta}}(B)$ so the solution u surely belongs to H^1.

That is $T : H^1_0(B) \to H^1_0(B)$ is an operator.

If we try to use Banach Fixed Point theorem, Theorem 18.1, we would need to show contraction, and consider

$$\Delta(Tu - Tv) = |Du|^\theta - |Dv|^\theta.$$

Multiplying with $Tu - Tv$, we get

$$\begin{aligned}\|D(Tu - Tv)\|^2_{L^2(B)} =& \|\left(|Du|^\theta - |Dv|^\theta\right)(Tu - Tv)\|_{L^1(B)}\\ \leq& \||Du|^\theta - |Dv|^\theta\|_{L^{\frac{2n}{n+2}}} \|Tu - Tv\|_{L^{\frac{2n}{n-2}}}\\ \lesssim& \||Du|^\theta - |Dv|^\theta\|_{L^{\frac{2n}{n+2}}} \|D(Tu - Tv)\|_{L^2(B)}.\end{aligned}$$

So we have shown

$$\|D(Tu - Tv)\|_{L^2(B)} \lesssim \||Du|^\theta - |Dv|^\theta\|_{L^{\frac{2n}{n+2}}}.$$

We have, Exercise 18.14,

$$\left||Du|^\theta - |Dv|^\theta\right| \lesssim |Du - Dv|^\theta,$$

and thus arrive at (last step is Hölder, since $\theta\frac{2n}{n+2} \leq 2$)

$$\|D(Tu - Tv)\|_{L^2(B)} \lesssim \|Du - Dv\|^\theta_{L^{\theta\frac{2n}{n+2}}} \lesssim \|Du - Dv\|^\theta_{L^2(B)},$$

By Poincaré we then have

$$\|Tu - Tv\|_{H^1(B)} \leq \|u - v\|^\theta_{H^1(B)}.$$

So if $\theta \in (0, 1)$, $T : H^1_0(B) \to H^1_0(B)$ is continuous (but not Lipschitz!), if $\theta = 1$ then T is indeed Lipschitz, but T is still not a contraction! So it is *not clear* how we would use the *Banach Fixed Point argument* from before.

But now we could check for the assumptions of Leray-Schauder, Corollary 18.10! First we need that $T : H^1_0(B) \to H^1_0(B)$ is *compact*.

Let $(u_k)_{k\in\mathbb{N}} \subset H^1_0(\Omega)$ be bounded, i.e.

$$\sup_{k\in\mathbb{N}} \|u_k\|_{H^1_0(\Omega)} < \infty.$$

We then have

$$\Delta T u_k = g + |Du_k|^\theta$$

Testing this equation with $T(u_k)$ we have

$$\|T(u_k)\|^2_{H^1} \leq \|DT(u_k)\|^2_{L^2(B)} \leq \|g\|_{L^2(B)}\|T(u_k)\|_{L^2} + \||Du_k|^\theta\|_{L^2(B)}\|T(u_k)\|_{L^2(B)}.$$

By Young's inequality $ab \leq \varepsilon a^2 + \frac{1}{\varepsilon}b^2$ we find for any $\varepsilon > 0$

$$\|T(u_k)\|^2_{H^1} \leq \|DT(u_k)\|^2_{L^2(B)} \leq \frac{C}{\varepsilon}\left(\|g\|^2_{L^2(B)} + \||Du_k|^\theta\|^2_{L^2(B)}\right) + \varepsilon C\|T(u_k)\|^2_{H^1(B)}.$$

Taking ε so that $\varepsilon C < \frac{1}{2}$ We can absorb to the left-hand side

$$\|T(u_k)\|^2_{H^1} \lesssim \|g\|^2_{L^2(B)} + \||Du_k|^\theta\|^2_{L^2(B)}$$

Since $\theta \in (0,1]$ we conclude that

$$\sup_k \|T(u_k)\|_{H^1} < \infty.$$

Then there exists a weakly convergent subsequence (not relabeled) $T(u_k) \xrightarrow{k\to\infty} w$ *weakly* in $H^1_0(B)$. We need strong convergence (of a subsequence) in $H^1_0(B)$! Taking a further subsequence (again not relabeled) we obtain from Rellich's theorem, Theorem 13.35, that $T(u_k) \xrightarrow{k\to\infty} w$ *strongly* in $L^2(B)$. As we have seen above, the right-hand side only depends on the L^2-norm of Tu_k, so we are going to use that to conclude convergence. We have

$$\Delta(Tu_k - Tu_\ell) = |Du_k|^\theta - |Du_\ell|^\theta$$

Testing yet again with $Tu_k - Tu_\ell$ we find

$$\|D(Tu_k - Tu_\ell)\|^2_{L^2(B)} \le \underbrace{\||Du_k|^\theta - |Du_\ell|^\theta\|_{L^2(B)}}_{\text{uniformly bounded in } k,\ell} \underbrace{\|Tu_k - Tu_\ell\|_{L^2(B)}}_{\xrightarrow{\ell,k\to\infty} 0}.$$

Thus $(DTu_k)_{k\in\mathbb{N}}$ is a Cauchy sequence and thus convergent, the limit is unique so Tu_k converges *strongly* to w in $H^1_0(B)$.

That is $\underline{T \text{ is compact}}$.

In order to apply Leray-Schauder, Corollary 18.10, we need to give a priori estimates of solutions to $\mu T u = u$, where $\mu \in [0,1]$ and $u \in H^1_0(B)$. If $\mu = 0$ there is nothing to show since then $u = 0$, so let $\mu > 0$. Then we have

$$\Delta T u = g + |Du|^\theta$$

and thus

$$\frac{1}{\mu}\Delta u = g + |Du|^\theta$$

i.e.

$$\Delta u = \mu g + \mu|Du|^\theta$$

What are we going to do? We are going to test with u.

$$\begin{aligned}\|u\|_{H^1_0(B)} \lesssim \|Du\|^2_{L^2(B)} \le & \mu\|g\|_{L^2}\|u\|_{L^2(B)} + \mu\||Du|^\theta\|_{L^2(B)}\,\|u\|_{L^2(B)} \\ \overset{\mu\le 1}{\lesssim} & C(\varepsilon)\|g\|^2_{L^2} + \varepsilon\|u\|^2_{L^2(B)} + \|Du\|^\theta_{L^{2\theta}(B)}\,\|u\|_{L^2(B)}\end{aligned}$$

So we have for small enough ε using that $\theta \le 1$,

$$\|u\|_{H^1_0(B)} \lesssim C(\varepsilon)\|g\|^2_{L^2} + \|u\|^{1+\theta}_{H^1_0(B)}$$

And here we need to use that $\theta < 1$, because then we have by Young's inequality for any $\delta > 0$

$$\|u\|^{1+\theta}_{H^1_0(B)} \le \delta\|u\|^2_{H^1(B)} + C(\delta).$$

Thus we have shown

$$\|u\|_{H^1_0(B)} \lesssim C(\varepsilon)\|g\|^2_{L^2} + C(\delta),$$

The right-hand side is clearly independent from μ and gives a uniform bound of solutions to $\mu T u = u$.

We finally can use Leray-Schauder, Corollary 18.10, to conclude there exists some $u \in H^1_0(B)$ such that $Tu = u$, and thus a solution to

$$\Delta u = g + |Du|^\theta$$

Exercise 18.14. *(1) Show that for each $\theta \in (0,1]$ there exists a constant $C = C(\theta) > 0$ such that*

$$|a^\theta - b^\theta| \leq C(\theta)|a-b|^\theta \quad \forall a, b \geq 0.$$

(2) Find a similar estimate if $\theta \geq 1$

19. Hilbert spaces

Hilbert spaces are very special spaces, so they deserve some extra attention.

Hilbert spaces are Banach spaces with a scalar product – if the space is not complete but has a scalar product . More precisely.

Definition 19.1. Let $(X, *, +)$ be a *linear space*, or *vector space*, $(X, *, +)$. A *scalar product* or *inner product* is a map $\langle \cdot, \cdot \rangle : X \to \mathbb{R}$ with the following properties

(1) $\langle x, y \rangle = \langle y, x \rangle \quad \forall x, y \in X$
(2) $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ if and only if $x = 0$
(3) $\langle \lambda x + \mu y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$ for all $\lambda, \mu \in \mathbb{R}$, $x, y, z \in X$.

A vector space equipped with a scalar product is called a *inner product space* or *pre-Hilbert space*, with the induced norm $\|x\| := \sqrt{\langle x, x \rangle}$. If it is complete it is called a Hilbert space.

Recall, Exercise 7.9, a norm $\|\cdot\|_X$ can be derived from a scalar product if and only if it satisfies the parallelogram identity

$$2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2 \quad \text{for all } x, y \in X.$$

We also recall Cauchy Schwarz inequality which holds whenever we have a scalar product (and the norm is the induced norm).

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

Examples for Hilbert spaces are ℓ^2, L^2, or $W^{k,2}$.

From Calculus we recall the formula (e.g. in $\mathbb{R}^3$)

$$\langle v, w \rangle = |v||w| \cos(\theta),$$

where θ is the angle between v and w. And this is really a feature of the Hilbert spaces, we can talk (reasonably) about angles – in particular we have orthogonality.

Definition 19.2. In a (Pre-)Hilbert space H we say

$$v \perp w \quad \text{if } \langle v, w\rangle = 0.$$

Example 19.3. Two functions $f, g \in L^2(\mathbb{R}^n)$ with $\operatorname{supp} f \cap \operatorname{supp} g = \emptyset$ are orthogonal in L^2. Indeed,

$$\langle f, g\rangle_{L^2} = \int fg = 0.$$

Exercise 19.4. *(1) Consider $f \in L^2(\mathbb{R}^n, \mathbb{R}^n)$*
(2) Let $u, v \in W^{1,2}(\mathbb{R}^n)$ and set

$$f := \nabla u = \begin{pmatrix} \partial_x u \\ \partial_y u \end{pmatrix}$$

$$g := \nabla^\perp v \equiv \begin{pmatrix} -\partial_y u \\ \partial_x u \end{pmatrix}$$

Show that $f \perp g$ in $L^2(\mathbb{R}^2, \mathbb{R}^2)$

Exercise 19.5. *Let H be a Pre-Hilbertspace. Show that if $x \perp y$ then*

$$\|x + y\|_H^2 = \|x\|^2 + \|y\|^2.$$

Orthogonality depends on the specific choice of scalar product. Even if two Hilbert-space norms are equivalent, two vectors being orthogonal with respect to one norm/scalar product may not be orthogonal with respect to the other.

Exercise 19.6. *(Cf. [StackExchange, b])*

1. *Let $\langle \cdot, \cdot\rangle$ be the usual scalar product for $\mathbb{R}^2$-vectors. Define a different scalar product $[\cdot, \cdot]$ with two vectors $v, w \in \mathbb{R}^2$ such that $v \perp w$ with respect to $\langle, \rangle$, but $v \not\perp w$ with respect to $[,]$. Show that the induced norms are equivalent (this is easy!),*
2. *Show that if V is equipped with two scalar produts $\langle, \rangle$ and $[,]$. Assume that we have*

$$\langle v, w\rangle = 0 \quad \Leftrightarrow [v, w] = 0 \quad \forall v, w \in V.$$

Show that then there exists $\lambda > 0$ such that

$$\langle v, w\rangle = \lambda[v, w] \quad \forall v, w \in V.$$

From Calculus we know that we can describe a plane P through the origin in $\mathbb{R}^3$ via its normal vector n, $P = n^\perp$. This construction works also in arbitrary spaces:

Definition 19.7. Let $v \in H$. Set

$$v^\perp := \{w \in H : \quad w \perp v\}.$$

More generally let $Y \subset H$ then its orthogonal complement

$$Y^\perp = \{v \in H : v \perp y \quad \forall y \in Y\}.$$

(in particular $\emptyset^\perp = H$)

Exercise 19.8. *Let $U \subset V \subset H$. Show that $V^\perp \subset U^\perp$.*

Lemma 19.9. *$Y^\perp$ is closed[43] linear subspace of H.*

Proof. Let $v, w \in Y^\perp$ and take $\lambda, \mu \in \mathbb{R}$. For any $y \in Y$ we have $v \perp y$, i.e. $\langle v, y\rangle = 0$ and $w \perp y$, i.e. $\langle w, y\rangle = 0$. Thus

$$\langle \lambda v + \mu w, y\rangle = \lambda\langle v, y\rangle + \mu\langle w, y\rangle = 0,$$

and thus $\lambda v + \mu w \perp y$. This holds for any $y \in Y$, so we have $\lambda v + \mu w \in Y^\perp$ – i.e. $Y^\perp$ is a linear subbspace of H.

The interesting part is that $Y^\perp$ is closed. This follows from the continuity of the scalar product. Let $v_k \in Y^\perp$ with $v = \lim_{k\to\infty} v_k$. Take any $y \in Y$, then

$$|\langle v, y\rangle| = |\langle v - v_k, y\rangle + 0| \leq \|v - v_k\|\|y\| \xrightarrow{k\to\infty} 0.$$

Thus $\langle v, y\rangle = 0$, which holds for any $y \in Y$, and thus $v \in Y^\perp$ making $Y^\perp$ closed. □

Lemma 19.10. *If $\overline{Y} = H$ then $Y^\perp = \{0\}$.*

Proof. Assume $y \in Y^\perp$. Since Y is dense in H there must be some sequence $Y \ni y_k \xrightarrow{k\to\infty} y$, and thus

$$0 \overset{y\in Y^\perp}{=} \langle y_k, y\rangle \xrightarrow{k\to\infty} \langle y, y\rangle = \|y\|^2.$$

That is $y = 0$, and thus $Y^\perp = \{0\}$. We can conclude. □

The opposite of Lemma 19.10 is clearly false. Take $Y = \{1\} \subset \mathbb{R}$ then $Y^\perp = 0$. However, if H is Hilbert and Y is a *linear space* then the converse is true, Proposition 19.12.

Lemma 19.11. *Let H be a Hilbert space and let $C \subset H$ be a closed convex nonempty set. Then for each $x \in H$ there exists exactly one $\bar{c} \in C$ with*

$$\|x - \bar{c}\| = \operatorname{dist}(x, C) = \inf_{\tilde{c}\in C} \|x - \tilde{c}\|.$$

If C is a linear space then $\bar{c}$ satisfies

$$x - \bar{c} \in C^\perp.$$

Proof. Fix $x \in H$. If $x \in C$ then we chose $\bar{c} := x$ which satisfies all the above assumptions.

In the more general case, since $d := \operatorname{dist}(x, C)$ is an infimum over a nonempty set there must be a sequence $c_i \in C$ such that

$$\|x - c_k\| \xrightarrow{k\to\infty} d.$$

We need to show that c_i converges (if we already had reflexivity of Hilbert spaces we could use weak convergence, however that prove is based on this lemma): For this what essentially amounts to quantifiable strict convexity of the unit disk.

[43]since we might be in infinite dimensions, linear subspaces might not be closed! Think of $C_c^\infty(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$ with the L^2-norm.

Consider c_k, c_ℓ for k and ℓ large. We could imagine that either $c_k \approx c_\ell$ or the convex combination $\frac{1}{2}c_k + \frac{1}{2}c_\ell$ should be "even closer". With the scalar product we can quantify this:

Consider $u, v \in C$ (in our imagination we should think $\|v-x\|$ and $\|u-x\|$ close to d, but this is not necessary for the argument):

$$\begin{aligned}\left\|x - \frac{v+u}{2}\right\|^2 =& \frac{1}{4}\|(x-v)+(x-u)\|^2\\
=& \frac{1}{4}\left(\|x-v\|^2 + \|x-u\|^2 + 2\langle x-v, x-u\rangle\right)\\
=& \frac{1}{4}\left(2\|x-v\|^2 + 2\|x-u\|^2 - \|x-v\|^2 - \|x-u\|^2 + 2\langle x-v, x-u\rangle\right)\\
=& \frac{1}{4}\left(2\|x-v\|^2 + 2\|x-u\|^2 - \|u-v\|^2\right)\end{aligned}$$

This is a very geometric inequality: if u and v are not very close to each other, then the convex combination $\frac{v+u}{2}$ will be closer to x than the worst distance of x to v or u.

In particular we have

$$\frac{1}{4}\|u-v\|^2 \leq \frac{1}{2}\|x-v\|^2 + \frac{1}{2}\|x-u\|^2 - \left\|x - \frac{v+u}{2}\right\|^2$$

Since C is convex, we have that $\frac{v+u}{2} \in C$ and thus $\left\|x - \frac{v+u}{2}\right\| \geq d$. We conclude

$$\frac{1}{4}\|u-v\|^2 \leq \frac{1}{2}\left(\|x-v\|^2 - d^2\right) + \frac{1}{2}\left(\|x-u\|^2 - d^2\right) \quad \forall u, v \in C. \tag{19.1}$$

If we set $u = c_k$ and $v = c_\ell$ and use that $\|c_k - x\|, \|c_\ell - x\| \xrightarrow{k,\ell\to\infty} d$ we conclude that $(c_\ell)_{\ell\in\mathbb{N}}$ is a Cauchy sequence, since H is Hilbert and C is closed we find a limit $\bar{c} \in C$. By continuity we conclude that

$$\|\bar{c} - x\| = dist(x, C).$$

We now prove uniqueness of $\bar{c}$ - and it relies on quantifiable strict convexity (19.1): If $\bar{c}$ and $\tilde{c}$ both satisfy $\|\bar{c} - x\| = \|\tilde{c} - x\| = dist(x, C)$ then by (19.1) we have that $\tilde{c} = \bar{c}$.

Lastly we need to show that (if C is a linear space!)

$$x - \bar{c} \in C^\perp.$$

Fix any $c \in C$, then $(1-t)\bar{c} + tc \in C$ for all $t \in [0,1]$. Set

$$f(t) := \|x - (1-t)\bar{c} - tc\|^2 - d^2$$

We then have $f(t) \geq 0$ for all $t \in [0,1]$ and $f(0) = 0$. We then know the sign of the one-sided derivative

$$\begin{aligned}
0 \leq \lim_{t\to 0^+} \frac{f(t)-f(0)}{t-0} &= \lim_{t\to 0^+} \frac{\|x-(1-t)\bar{c}-tc\|^2 - \|x-\bar{c}\|^2}{t} \\
&= \lim_{t\to 0^+} \frac{\|x-\bar{c}-t(c-\bar{c})\|^2 - \|x-\bar{c}\|^2}{t} \\
&= \lim_{t\to 0^+} \frac{\|x-\bar{c}\|^2 + t^2\|c-\bar{c}\|^2 - 2t\langle c-\bar{c}, x-\bar{c}\rangle - \|x-\bar{c}\|^2}{t} \\
&= 2\langle c-\bar{c}, x-\bar{c}\rangle
\end{aligned}$$

So we have

$$0 \leq \langle c-\bar{c}, x-\bar{c}\rangle \quad \forall c \in C.$$

If C is a linear space, we can take $c = \bar{c} + \tilde{c}$ and have

$$0 \leq \langle \tilde{c}, x-\bar{c}\rangle \quad \forall \tilde{c} \in C.$$

We can also take $-\tilde{c}$ and thus have

$$0 \leq \langle -\tilde{c}, x-\bar{c}\rangle \quad \forall \tilde{c} \in C.$$

which implies

$$0 \geq \langle \tilde{c}, x-\bar{c}\rangle \quad \forall \tilde{c} \in C.$$

Together we have shown

$$0 = \langle \tilde{c}, x-\bar{c}\rangle \quad \forall \tilde{c} \in C$$

and thus can conclude. □

Proposition 19.12. *H Hilbert and Y linear space. Then $Y^\perp = \{0\}$ means $\overline{Y} = H$.*

Proof. Let $z \in H$. Since $\overline{Y}$ is closed and convex, by Lemma 19.11 there must be some $y \in \overline{Y}$ with $\operatorname{dist}(z, \bar{Y}) = \|z-y\|$ – and we have

$$y - z \in Y^\perp.$$

Since $Y^\perp = \{0\}$ we find that $y = z$ and thus $z \in \bar{Y}$. That is $\bar{Y} = H$. □

Lemma 19.13. *Let H be a Hilbert space and Y a closed linear space. Then $H = Y \oplus Y^\perp$ in the sense that for any $x \in H$ there exists a unique $y \in Y$ and a unique $\tilde{y} \in Y^\perp$ such that*

$$x = y + \tilde{y}$$

and we have

$$\|x\|^2 = \|y\|^2 + \|y^\perp\|^2.$$

Proof. First we discuss uniqueness. If $y_1, y_2 \in Y$ and $\tilde{y}_1, \tilde{y}_2 \in Y^\perp$ such that

$$x = y_1 + \tilde{y}_1 = y_2 + \tilde{y}_2.$$

Then

$$y_1 + \tilde{y}_1 = y_2 + \tilde{y}_2,$$

then we have

$$Y \ni y_1 - y_2 = \tilde{y}_2 - \tilde{y}_1 \in Y^\perp$$

Thus $y_1 - y_2$ and $\tilde{y}_2 - \tilde{y}_1 \in Y^\perp \cap Y$. But if $y \in Y \setminus \{0\}$ then $\langle y, y\rangle = \|y\|^2 > 0$ so $y \not\in Y^\perp$. Thus $Y \cap Y^\perp = \{0\}$. This implies $y_1 = y_2$ and $\tilde{y}_1 = \tilde{y}_2$, so uniqueness is established.

Now let $x \in H$. Since Y is a closed linear space, by Lemma 19.11 there exists a unique $y \in Y$ such that

$$\|y - x\| = \mathrm{dist}\,(x, Y).$$

and we have

$$\tilde{y} := y - x \in Y^\perp.$$

We can conclude. □

Lemma 19.14. *Let H be a Hilbert-space. Consider $Y^{\perp\perp} = \left(Y^\perp\right)^\perp$. Then we have*

$$Y^{\perp\perp} = \overline{\mathrm{span} Y}.$$

In particular if Y is a closed linear subspace we have $Y^{\perp\perp} = Y$.

Proof. Clearly we have $Y^{\perp\perp} \supset Y$.

Set $W := \overline{\mathrm{span} Y}$. By Exercise 19.8 we have $W^\perp \subset Y^\perp$, and thus

$$W^{\perp\perp} \supset Y^{\perp\perp}.$$

Also observe that for any $v \in Y^\perp$ and any $y \in \mathrm{span} Y$ we have $y \perp v$ by lienarity of the scalar product. Thus $\mathrm{span} Y \subset Y^{\perp\perp}$, and with Lemma 19.9 we find that

$$W^{\perp\perp} \supset Y^{\perp\perp} \supset \overline{\mathrm{span} Y} = W.$$

It suffices thus to show that $W^{\perp\perp} \subset W$, then the previous set inequalities are equalities and we can conclude.

Take now any $y \in W^{\perp\perp}$. Since $W^{\perp\perp}$ is a closed linear subspace of a Hilbert space, $W^{\perp\perp}$ is itself a Hilbert space. So we can apply Lemma 19.11 and find $w \in W \subset W^{\perp\perp}$ with $w = \mathrm{dist}\,(y, W)$, and then we have $v := y - w \in W^\perp \cap W^{\perp\perp}$. But then $\langle v, v\rangle = 0$, which means $v = 0$ which means $y = w \in W$. We have shown that $W^{\perp\perp} \subset W$ as requested and we can conclude. □

Theorem 19.15. *Let H be a Hilbert space, then the following are equivalent*

(1) H is separable (i.e. there exists a countable dense set)

(2) H has a countable orthonormal basis, i.e. there exists a sequence $(o_k)_{k=1}^{\dim H}$ for each $x \in H$ we have

$$x = \sum_{k=1}^{\dim H} o_k \langle x, o_k\rangle$$

where the convergence of the (possibly infinite) sum is in H.

We only prove (1) $\Rightarrow$ (2), the other direction follows from the exercise Exercise 19.17.

Proof of Theorem 19.15: (1) $\Rightarrow$ (2). Let $Q \subset H$ be a countable dense set. In particular

$$\overline{\operatorname{span} Q} = H.$$

Where span denotes all the finite linear combinations of Q.

We can now pass to a smaller set $\tilde{Q}$ such that still

$$\overline{\operatorname{span} \tilde{Q}} = H,$$

but each finite subset of $\tilde{Q}$ is linearly independent (simply consecutively remove linearly dependent vectors, which we can do since Q is countable). Let N be the number of elements in $\tilde{Q}$ (clearly $N \leq \dim H$).

We write

$$\tilde{Q} = (v_1, v_2, \ldots).$$

Since each finite subset of $\tilde{Q}$ is linear independent we know $|v_1| \neq 0$, so we may set

$$o_1 := \frac{v_1}{|v_1|}.$$

Now we define o_i, $i = 1, \ldots, N$ inductively by

$$\tilde{o}_{i+1} := v_{i+1} - \sum_{i=1}^{n} o_i \langle o_i, v_{i+1} \rangle$$

and set

$$o_{i+1} := \frac{o_i}{|o_{i+1}|}.$$

This process is called *Gram-Schmidt*-orthogonalization of linearly independent vectors. For this we observe that by linear independence $\tilde{o}_{i+1} \neq 0$ and $\tilde{o}_{i+1}$ is linear independent of $(o_k)_{k=1}^{i}$.

By induction we conclude that $(o_i)_{i=1}^{N}$ is a (finite or infinite) sequence of mutually orthogonal vectors, all satisfying $\|o_i\| = 1$.

On the other hand we still have

$$\overline{\operatorname{span}\{o_i : \quad i = 1, \ldots\}} = \overline{\operatorname{span} \tilde{Q}} = H,$$

since o_i are each (finite) linear combinations of elements of $\tilde{Q}$.

We claim that $(o_k)_{k=1}^{N}$ is a basis as in (2). So let $x \in H$.

Consider for $\ell = 1, \ldots, N$,

$$y_\ell := x - \sum_{k=1}^{\ell} o_k \langle x, o_k \rangle.$$

Then we have

$$y_\ell \perp \operatorname{span}\{o_1, \ldots, o_\ell\} \quad \forall \ell.$$

Thus,

$$\|x\| = \|y_\ell\|^2 + \sum_{k=1}^{\ell} |\langle x, o_k\rangle|^2$$

In particular we get

$$\sum_{k=1}^{N} |\langle x, o_k\rangle|^2 < \infty,$$

and thus $y_N \in H$, defined as

$$y_N := x - \sum_{k=1}^{N} o_k \langle x, o_k\rangle$$

exists. If $y_N = 0$ we are done. Assume $y_N \neq 0$ then

$$0 \neq y_N \perp o_k \quad \forall k = 1, \ldots, N$$

In particular, for any $z \in \operatorname{span}\{o_1, \ldots, o_N\}$

$$\|y_N - z\|^2 = \|y_N\| + \|z\|^2 \geq \|y_N\|^2 > 0.$$

That is $\operatorname{dist}(y_N, \operatorname{span}\{o_1, \ldots, o_N\}) > 0$ which implies

$$y_N \notin \operatorname{span}\{o_1, \ldots, o_N\},$$

which is a contradiction to the construction of $(o_i)_i$.

Thus $y_N = 0$ and we can conclude. □

Exercise 19.16. *Assume that H is a Hilbert space, $(o_k)_{k\in\mathbb{N}}$ with $\|o_k\| = 1$ and $o_k \perp o_j$ for $k \neq j$. Assume that $x \in X$ satisfies*

$$x = \sum_{k=1}^{\infty} \lambda_k o_k \equiv \lim_{N\to\infty} \sum_{k=1}^{N} \lambda_k o_k,$$

(convergence in H).

Then

$$\sum_{k=1}^{\infty} |\lambda_k|^2 = \|x\|^2.$$

Exercise 19.17. *Show that every seperable Hilbert space is isometric isomorphic to $\ell^2(\mathbb{N})$.*

Hint: Use the representation (2) in Theorem 19.15 and show with the help of Exercise 19.16 that

$$x \mapsto (\langle x, o_k\rangle)_{k\in\mathbb{N}}$$

is an isometric isomorphism.

19.1. Riesz Representation for Hilbert spaces and reflexivity.

Theorem 19.18 (Riesz representation theorem)**.** *Let H be a Hilbert space. Let $x^* \in H^*$, then there exists exactly one $\bar{x} \in H$ such that*

$$x^*[y] = \langle \bar{x}, y\rangle \quad \forall y \in Y.$$

Moreover, $\|x^*\|_{H^*} = \|\bar{x}\|_H$.

Proof. If $x^* = 0$ there is nothing to show, we just take $\bar{x} = 0$.

So without loss of generality we may assume

$$\|x^*\|_{H^*} = 1.$$

We take any sequence $(x_k)_{k\in\mathbb{N}}$, $\|x_k\| = 1$ such that

$$x^*(x_k) \xrightarrow{k\to\infty} \|x^*\| = 1.$$

We now essentially use strict convexity of the unit ball of H: $\frac{x_k+x_\ell}{2}$ is strictly inside that unit ball. Using the scalar product structure we can quantify this (the following argument should remind you of the argument used in Lemma 19.11):

$$\begin{aligned}
\left\|\frac{x_k+x_\ell}{2}\right\|_H^2 =&\frac{1}{4}\left(\|x_k\|^2 + \|x_\ell\|^2 + 2\langle x_k, x_\ell\rangle\right)\\
=&\frac{1}{4}\left(2 + 2\langle x_k, x_\ell\rangle\right)\\
=&\frac{1}{4}\left(4 - 2 + 2\langle x_k, x_\ell\rangle\right)\\
=&\frac{1}{4}\left(4 - \|x_k\|^2 - \|x_\ell\|^2 + 2\langle x_k, x_\ell\rangle\right)\\
=&\frac{1}{4}\left(4 - \|x_k - x_\ell\|^2\right)\\
=&1 - \left\|\frac{x_k - x_\ell}{2}\right\|^2
\end{aligned} \tag{19.2}$$

Let $\varepsilon \in (0,1)$, take k,ℓ large enough so that $x^*(x_k) > 1-\varepsilon$ and $x^*(x_\ell) > 1-\varepsilon$. Then

$$\begin{aligned}
1-\varepsilon \leq&\frac{1}{2}\left(x^*(x_k) + x^*(x_\ell)\right)\\
=&x^*\left(\frac{x_k+x_\ell}{2}\right)\\
\precsim&\underbrace{\|x^*\|}_{=1}\left\|\frac{x_k+x_\ell}{2}\right\|
\end{aligned}$$

Squaring this inequality we have

$$1+\varepsilon^2 - 2\varepsilon \leq 1 - \left\|\frac{x_k - x_\ell}{2}\right\|^2$$

and thus
$$\left\|\frac{x_k - x_\ell}{2}\right\|^2 \leq 2\varepsilon - \varepsilon^2 \overset{\varepsilon<\varepsilon^2}{\leq} \varepsilon.$$
This holds for any $\varepsilon \in (0,1)$ for all large k and ℓ, so we have shown that $(x_k)_k$ is Cauchy, so since we are in a Hilbert space there exists a limit $\bar{x} = \lim_{k\to\infty} x_k$. By continuity we have $\|\bar{x}\| = 1$ and
$$x^*[\bar{x}] = 1.$$
Assume there is any other $\tilde{x}$ with $\|\tilde{x}\| = 1$ such that $x^*(\tilde{x}) = 1$. Set $w := \frac{\tilde{x}+\bar{x}}{2}$. We have
$$x^*(w) = 1$$

and in particular
$$x^*\left(\frac{w}{\|w\|}\right) = \frac{1}{\|w\|}$$
If $\tilde{x} \neq \bar{x}$ we have $\|w\| < 1$ ((19.2) – strict convexity of the unit ball), and thus we have
$$x^*\left(\frac{w}{\|w\|}\right) > 1,$$
which is a contradiction to $\|x^*\| = 1$. So $\tilde{x} = \bar{x}$, i.e. $\bar{x}$ is the unique vector such that
$$\|\bar{x}\|_H = 1 \quad \text{and} \quad x^*[\bar{x}] = \|x^*\|.$$
(so far the above argument is based only on the (quantifiable) strict convexity of the unit ball in the space X)

We now claim that
$$x^*[y] = \langle \bar{x}, y\rangle \quad \forall y \in H.$$
If $y = \lambda\bar{x}$ this is clear, because
$$x^*[\lambda\bar{x}] = \lambda = \langle \bar{x}, \lambda\bar{x}\rangle.$$
Now let $y \in \bar{x}^\perp$. We claim that $x^*(y) = 0$. W.l.o.g. we may assume that $\|y\| = 1$. For any $t \in \mathbb{R}$ we then have by orthogonality
$$\|\bar{x} + ty\|^2 = \|\bar{x}\|^2 + t^2\|y\|^2 = 1 + t^2.$$
Set
$$z(t) := \frac{\bar{x} + ty}{1+t^2}.$$
Then $\|z(t)\| = 1$ for all t, $z(0) = \bar{x}$, and we have
$$y = \frac{d}{dt}\Big|_{t=0} z(t) \equiv \lim_{t\to 0} \frac{\frac{x+ty}{1+t^2} - x}{t}.$$
Setting now
$$f(t) := x^*(z(t))$$

we have that $f(0) = x^*(\bar{x}) = 1$ and $f(t) = x^*(z(t)) \leq \|x^*\| = 1$ for all t. That is f attains its maximum at $t = 0$ and $f'(0) = 0$ (checking the differentiability of f is an exercise). But then

$$x^*(y) = f'(0) = 0.$$

Thus, we conclude that $x^*(y) = 0$ for all $y \in Y^\perp$.

Let now $y \in H$. In view of Lemma 19.13, we can uniquely split y into

$$y = \lambda\bar{x} + \tilde{y},$$

where $\tilde{y} \in \bar{x}^\perp$. We then have

$$x^*(y) = \lambda \underbrace{x^*(\bar{x})}_{=1} + \underbrace{x^*(\tilde{y})}_{=0} = \lambda = \langle \bar{x}, \lambda\bar{x}\rangle + \langle y, \bar{x}\rangle = \langle y, \bar{x}\rangle.$$

We can conclude. □

With Theorem 19.18 the structure of the dual space of a Hilbert space is very similar to the Hilbert space itself. Firstly we observe

Corollary 19.19. *Let H be a Hilbert space. Then H and H^* are isomorphic under the isomorphism*

$$H \ni x \mapsto x^*,$$

where for fixed $x \in X$ we set

$$x^*[y] := \langle x, y\rangle \quad y \in Y.$$

Exercise 19.20. *Prove Corollary 19.19.*

We can define a scalar product for H^*:

For $x^*, y^* \in H^*$ let x, y be the (unique) representatives of x^* and y^*, respectively, from Theorem 19.18. Then we set

$$\langle x^*, y^*\rangle_{H^*} := \langle x, y\rangle_H$$

Exercise 19.21. *Show that $\langle \cdot, \cdot\rangle_{H^*}$ defined as above*

(1) defines indeed a scalar product in H^* and
(2) induces the norm $\|\cdot\|_{H^*}$.
(3) Conclude that H^* is a Hilbert space.

This is what it means when we sometimes say $H^* = H$ for a Hilbert space.

If $H^* = H$ then certainly $H^{**} = H^* = H$ – so we should have reflexivity? Yes, indeed.

Corollary 19.22. *Let H be a Hilbert space, then H is reflexive.*

Proof. Recall, cf. Section 11, that the canonical isomorphism

$$J_H : H \mapsto H^{**}$$

is given

$$J_H : H \ni x \mapsto x^{**} \in H^{**}.$$

Here for fixed x the functional $x^{**} \in H^{**}$ is given by

$$x^{**}(y^*) := y^*(x) \quad \text{for } y^* \in H^*.$$

J_H is always injective and an isometry from H to H^{**}. H is called reflexive if J_H is moreover *surjective*.

So let $x^{**} \in H^{**}$. Since by Exercise 19.21 H^* is a Hilbert space, we can apply Riesz representation theorem, Theorem 19.18, and find some $x^* \in H^*$ such that

$$x^{**}[y^*] = \langle x^*, y^* \rangle_{H^*} \quad \forall y^* \in H^*. \tag{19.3}$$

Again applying Theorem 19.18 to H, we find the representative $x \in H$ such that

$$x^*[y] = \langle x, y \rangle_H \quad \forall y \in H.$$

Clearly x is a plausible candidate to satisfy $J_X x = x^{**}$. To see this is correct, we need to show

$$x^{**}[y^*] = J_X x[y^*] \equiv y^*[x] \quad \forall y^* \in H^*. \tag{19.4}$$

Fix any $y^* \in H^*$. Again by Theorem 19.18 there exists $y \in H$ such that

$$y^*[z] = \langle y, z \rangle \quad \forall z \in H. \tag{19.5}$$

In particular we have

$$y^*[x] \overset{(19.5)}{=} \langle x, y \rangle_H \overset{def}{\equiv} \langle x^*, y^* \rangle_{H^*} \overset{(19.3)}{=} x^{**}[y^*].$$

This shows (19.4) and we can conclude. □

19.2. **Lax-Milgram Theorem and application to existence theory.** Another important application of the Hilbert-space Riesz representation theorem is

Theorem 19.23 (Lax-Milgram)**.** *Let H be a Hilbert space, $A : H \times H \to \mathbb{R}$ be a bilinear and continuous map*[44]*, such that for some $\Lambda > 0$*

$$|A(x, y)| \leq \Lambda \|x\|_H \, \|y\|_H \quad \forall x, y \in H,$$

and for some $\lambda > 0$

$$|A(x, x)| > \lambda \|x\|_H^2 \quad \forall x \in H \tag{19.6}$$

Then there exists a bounded linear isomorphism $T : H \to H$ such that

$$A(x, y) = \langle Tx, y \rangle \quad \forall x, y \in H,$$

and we have

$$\|T\|_{\mathcal{L}(H,H)} \leq \Lambda, \quad \|T^{-1}\|_{\mathcal{L}(H,H)} \leq \lambda^{-1}$$

[44] may be not symmetric!

Proof. Fix $x \in H$ then the map

$$H \ni y \mapsto A(x,y) \in \mathbb{R}$$

is linear and bounded – thus it belongs to H^*. Consequently, by the Riesz Representation Theorem, Theorem 19.18, there exists exactly on element of H, we shall call it suggestively Tx (because it depends on x) such that

$$A(x,y) = \langle Tx, y\rangle \quad \forall x,y \in H.$$

It seems like this is it, but obverse: we don't know much about T – we still need to establish it is a bounded linear isomorphism $H \to H$.

As for linearity: Let $x_1, x_2 \in H$ and $\mu_1, \mu_2 \in \mathbb{R}$. We then have by definition of T

$$A(x_1,y) = \langle T(x_1), y\rangle, \quad A(x_2,y) = \langle T(x_2), y\rangle \quad \forall y \in H.$$

Then we have for any $y \in H$,

$$\begin{aligned}&\langle T(\mu_1 x_1 + \mu_2 x_2), y\rangle\\ =&A(\mu_1 x_1 + \mu_2 x_2, y)\\ =&\mu_1 A(x_1,y) + \mu_2 A(x_2,y)\\ =&\mu_1 \langle Tx_1, y\rangle + \mu_2\langle Tx_2, y\rangle\end{aligned}$$

Thus

$$\langle T(\mu_1 x_1 + \mu_2 x_2) - \mu_1 Tx_1 - \mu_2 Tx_2, y\rangle = 0 \quad \forall y \in H.$$

Taking $y = T(\mu_1 x_1 + \mu_2 x_2) - \mu_1 Tx_1 - \mu_2 Tx_2$ we find that

$$\|T(\mu_1 x_1 + \mu_2 x_2) - \mu_1 Tx_1 - \mu_2 Tx_2\|_H^2 = 0$$

and thus

$$T(\mu_1 x_1 + \mu_2 x_2) - \mu_1 Tx_1 - \mu_2 Tx_2 = 0$$

i.e. we have shown

$$T(\mu_1 x_1 + \mu_2 x_2) = \mu_1 Tx_1 + \mu_2 Tx_2 = 0.$$

That is, we have established linearity of T (we haven't used (19.6) for this).

Now we show that T is bounded. Observe that

$$|\langle Tx, y\rangle| = |A(x,y)| \leq \Lambda \|x\| \|y\|$$

Let now $y := Tx$ then we have

$$\|Tx\|^2 = |\langle Tx, Tx\rangle| \leq \Lambda \|x\| \|Tx\|$$

Dividing by $\|Tx\|$ we find that $\|Tx\| \leq \Lambda \|x\|$ which means that $\|T\|_{\mathcal{L}(H,H)} \leq \Lambda$. Thus $T : H \to H$ is a linear bounded operator.

Next we show injectivity. Let $x \in H$. Then we have

$$A(x,y) = \langle Tx, y\rangle \quad \forall y \in H.$$

Set $y := x$, using (19.6) we find with Cauchy-Schwartz,

$$\lambda \|x\|^2 \leq \langle Tx, x\rangle \leq \|Tx\| \|x\|,$$

dividing this inequality by $\|x\|$ we have shown

$$\lambda\|x\| \leq \|Tx\|,$$

which implies injectivity.

We can also use this to show closedness of TH. Indeed, let $x_k \in H$ and assume that $Tx_k \xrightarrow{k\to\infty} z \in H$. We then have

$$\lambda\|x_k - x_\ell\| \leq \|Tx_k - Tx_\ell\|,$$

and thus since $(Tx_k)_{k\in\mathbb{N}}$ is a Cauchy sequence, so is x_k (here we use that $\lambda > 0$). Since H is a Hilbert space there exists $\bar{x} = \lim_{k\to\infty} x_k$. By continuity we have

$$T\bar{x} = \lim_{k\to\infty} Tx_k = z.$$

That is $\bar{x} = T^{-1}z$. We conclude that $TH \subset H$ is closed.

Now we want to show that $TH = H$ (i.e. T is surjective). In view of Proposition 19.12 and closedness, $\overline{TH} = TH$, it suffices to show that $(TH)^\perp = \{0\}$. Assume $z \in (TH)^\perp$. Then we have

$$\langle Tx, z\rangle = 0 \quad \forall x \in H$$

By the definition of A we conclude that

$$A(x, z) = 0 \quad \forall x \in H.$$

But then we can take $x = z$ and have

$$\lambda\|z\|^2 \leq A(z, z) = 0,$$

that is (since $\lambda > 0$) we have $z = 0$. Thus $TH^\perp = \{0\}$. We conclude that T is linear a bijection with bounded $\|T\|$ and $\|T^{-1}\|$. That is, T is a linear bounded isomorphism. □

The following very useful statement, a consequence of Theorem 19.23, is often referred to as *Lax-Milgram Theorem* as well. It has more the feeling of *solving* an equation (and we will see this is exactly what it does in PDE).

Corollary 19.24 (Lax-Milgram)**.** *Let H be a Hilbert space, $A : H \times H \to \mathbb{R}$ be a bilinear and continuous map*[45]*, such that for some $\Lambda > 0$*

$$|A(x, y)| \leq \Lambda\|x\|_H \, \|y\|_H$$

and for some $\lambda > 0$

$$|A(x, x)| > \lambda\|x\|_H^2.$$

Assume $z^* \in H^*$. Then there exists exactly one $\bar{x} \in H$ such that

$$A(\bar{x}, y) = z^*(y) \quad \forall y \in H$$

and $\bar{x}$ satisfies

$$\|\bar{x}\|_H \leq \lambda^{-1}\|z^*\|_{H^*}.$$

[45] may be nonsymemtric!

Proof. In view of Theorem 19.23 we find $T : H \to H$ a linear bounded isomorphism such that

$$A(x, y) = \langle Tx, y\rangle \quad \forall x, y \in H.$$

Now let $z^* \in H^*$. By Riesz Represenation Theorem, Theorem 19.18, there exists an $\bar{z} \in H$ such that

$$z^*[y] = \langle \bar{z}, y\rangle.$$

Set $\bar{x} := T^{-1}\bar{z}$. Then we have the estimate

$$\|\bar{x}\| \leq \|T^{-1}\| \, \|\bar{z}\| \leq \lambda^{-1}\|\bar{z}\|.$$

Morever,

$$A(\bar{x}, y) = \langle T\bar{x}, y\rangle = \langle TT^{-1}\bar{z}, y\rangle = \langle \bar{z}, y\rangle = z^*[y] \quad \forall y \in H.$$

We can conclude. □

The typical application of Lax-Milgram is as follows.

Example 19.25. We want to solve the PDE

$$-\mathrm{div}\,(A\nabla u) = f \quad \text{in } \Omega$$

where $u \in W_0^{1,2}(\Omega)$. We know this is equivalent to solving

$$\int_\Omega A_{\alpha\beta}\partial_\beta u \cdot \partial_\alpha v = f[v].$$

Let $f \in (W_0^{1,2}(\Omega))^*$ and assume $A : \Omega \to \mathbb{R}^{n\times n}$ is measurable, *uniformly elliptic*, i.e. for some constant $\lambda > 0$

$$\langle Av, v\rangle \geq \lambda |v|^2 \quad \text{in } \Omega$$

and A is bounded $\sup_\Omega |A| \leq \Lambda$.

We can set

$$A(u, v) := \int_\Omega A_{\alpha\beta}\partial_\beta u \cdot \partial_\alpha v.$$

Then by ellipticity,

$$A(u, u) \geq \lambda \int_\Omega |\nabla u|^2 \gtrsim \|u\|_{W^{1,2}(\Omega)},$$

where the last inequality is Poincaré inequality.

And by boundedness,

$$|A(u, v)| \lesssim \Lambda \|\nabla u\|_{L^2(\Omega)} \, \|\nabla v\|_{L^2(\Omega)} \leq \Lambda \|u\|_{W^{1,2}(\Omega)} \, \|\nabla v\|_{W^{1,2}(\Omega)}.$$

Thus the conditions for Lax-Milgram Theorem 19.23 are satisfied, and we find that there exists a unique solution $u \in W_0^{1,2}(\Omega)$ to

$$\int_\Omega A_{\alpha\beta}\partial_\beta u \cdot \partial_\alpha v = f[v].$$

20. Open Mapping, Inverse Mapping, Closed Graph Theorem

We now begin to discuss some important results from classical Functional analysis, and for this review briefly the main definitions

- Normed spaces, Banach spaces, Hilbert spaces, Section 7
- linear bounded map, Section 8

20.1. Weak Baire Category Theorem.

These sections are inspired by [Megginson, 1998] and [StackExchange, a].

We want to avoid the full notion of Baire category, instead we prove a weak version of the Baire Category theorem, that implies as important consequences Zabreiko's Lemma, Lemma 20.8, which in turn implies such important theorems as open mapping theorem etc.

Definition 20.1. Let X be a vector space. A set $A \subset X$ is called *absorbing* if for each $x \in X$ there exists a number $s_x > 0$ such that

$$\frac{1}{t}x \in A \quad \forall t > s_x.$$

Clearly if X is a normed space and A is any open set containing the origin, then A is absorbing. The converse may not be true in general, see Exercise 20.3 below, but we have the following (which is a weak version of the *Baire cateogory theorem*)

Theorem 20.2. *Let X be a Banach space and $C \subset X$ be a closed, convex, and absorbing set. Then there exists an open set $U \subset C$ containing the origin $0 \in U$.*

Proof. Let

$$D := \{x \in X : x \in C, \quad -x \in C\} \subset C.$$

Observe that D is closed and convex. Consider the interior D^o of D

$$D^o = \{x \in D : \text{such that } \exists r > 0 \text{ with } B(x,r) \subset D\}.$$

Clearly, D^o is an open set. In general it could be empty, but if we can show that $D^o \neq \emptyset$, then we can set

$$U := \{x : \quad x = \frac{1}{2}a - \frac{1}{2}b \quad a, b \in D^o\},$$

which is an open set. Moreover if $x \in D^o$ then $-x \in D^o$, so $D^o \neq \emptyset$ implies $0 \in U$. Lastly, if $a \in D^o$ and $b \in D^o$ then $-b \in D^o$ and thus by convexity of D,

$$\frac{1}{2}a - \frac{1}{2}b = \frac{1}{2}a + \frac{1}{2}(-b) \in D,$$

and thus $U \subset D \subset C$. Thus U is the open set we were looking for. We just need to show that D^o is nonempty.

Assume to the contrary that $D^o = \emptyset$.

We are going consider for $n \in \mathbb{N}$,

$$nD := \{nx : \quad x \in D\}.$$

If $D^o = \emptyset$ then $(nD)^o = \emptyset$ for all $n \in \mathbb{N}$.

This implies that for each $n \in \mathbb{N}$ the set $X \setminus nD$ is dense in X. Moreover $X \setminus nD$ is open (since nD is closed).

Since $X \setminus D$ is open (and since its dense $X \setminus D \neq \emptyset$), there must be $x_1 \in X$ and $r_1 \in (0,1)$ such that $\overline{B(x_1, r_1)} \subset X \setminus D$.

Since $X \setminus 2D$ is open, we have that $B(x_1, r_1) \setminus 2D = (X \setminus 2D) \cap B(x_1, r_1)$ is open. Moreover, since $X \setminus 2D$ is dense in X, $B(x_1, r_1) \setminus 2D$ is nonempty. So there must be $x_2 \in X$, $r_2 \in (0, 1/2)$ such that $\overline{B(x_2, r_2)} \subset B(x_1, r_1) \setminus 2D$.

We continue like this and find a sequence of points $(x_i)_{i=1}^\infty$ and radii $r_i \in (0, \frac{1}{i})$ such that

$$B(x_i, r_i) \subset B(x_{i-1}, r_{i-1}) \setminus iD \neq \emptyset.$$

We then have that

$$|x_i - x_\ell| \leq 2 \max\{\frac{1}{i}, \frac{1}{\ell}\},$$

that is $(x_i)_{i \in \mathbb{N}}$ is a Cauchy sequence and thus $\bar{x} := \lim_{i \to \infty} x_i \in X$ (since X is a Banach space).

We then have for any $i \in \mathbb{N}$

$$\bar{x} \in \overline{B(x_i, r_i)} \subset X \setminus iD.$$

That is, $\frac{1}{i}\bar{x} \notin D$ for all $i \in \mathbb{N}$.

On the other hand, since C is absorbing, there must be some $\ell \in \mathbb{N}$ such that $\frac{1}{\ell}\bar{x} \in C$ and $\frac{1}{\ell}(-\bar{x}) \in C$, and thus $\frac{1}{\ell}\bar{x} \in D$, contradicting the previous statement. Thus $X \setminus iD$ cannot be dense for all i, meaning in particular D cannot have nonempty interior.

□

Exercise 20.3. *Let $X = (\ell^1(\mathbb{N}), \|\cdot\|_{\ell^\infty})$.*

(1) Show that X is not a Banach space.
(2) Consider the set

$$C := \{(a_n)_{n \in \mathbb{N}} \in X : \quad \|a_n\|_{\ell^1}\} \leq 1\}$$

Show that C is closed, convex and absorbing. However show that C has empty interior.

20.2. **Zabreiko's Lemma.** For us the important consequence of Theorem 20.2 (which is a weak version of Baire's theorem that we will not discuss here) is *Zabreiko's Lemma*, Lemma 20.8. It implies many of the results that are usually proven by Baire's category theorem.

Definition 20.4. A seminorm on a vector space X is a map $p : X \to \mathbb{R}$ such that

(1) $p(\lambda x) = |\lambda| p(x)$ for all $x \in X$, $\lambda \in \mathbb{R}$
(2) $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$

(observe this is lightly different from sublinearity in Hahn-Banach, Theorem 10.2, which assumes the multiplicity only for $\lambda \ge 0$)

Exercise 20.5. *Show $p(0) = 0$ for any seminorm p on a vector space X*

Exercise 20.6. *Show the reverse triangle inequality holds for any seminorm p on a vector space X*

$$|p(x) - p(y)| \le p(x - y) \quad \forall x, y \in X.$$

Exercise 20.7. *Let p be any seminorm on a vector space X. Show that if p is continuous at 0 then p is continuous in any point $x \in X$.*

Lemma 20.8 (Zabreiko's Lemma)**.** *Assume $p : X \to \mathbb{R}$ is a nonnegative seminorm on a Banach space X that is moreover* countably subadditive, *namely*

$$p(\sum_{n=1}^{\infty} x_n) \le \sum_n p(x_n)$$

whenever $(x_n)_{n\in\mathbb{N}} \subset X$ such that

$$\sum_{n=1}^{\infty} x_n := \lim_{N\to\infty} \underbrace{\sum_{n=1}^{N} x_n}_{\in X} \quad \text{in } X$$

Then p is continuous.

Exercise 20.9. *Let $(X, \|\cdot\|)$ be a normed space.*

(1) Show that $p(x) := \|x\|$ is a countably subadditive seminorm.
(2) Let $p : X \to \mathbb{R}$ be any nonnegative seminorm on X. Show that if p is continuous then p is countably subadditive.

Proof of Lemma 20.8. By Exercise 20.7 <u>we need to show that p is continuous in 0</u>.

Set

$$G := \{x \in X : p(x) < 1\}$$

Clearly $0 \in G$ and if $x \in X$ then for $s_x := 2\|x\|$ we have $\frac{1}{t}x \in G$ for any $t > s_x$ since p is a seminorm. That is G is absorbing.

G is also convex, indeed if $x, y \in G$ and $\lambda \in (0,1)$ then

$$p(\lambda x + (1-\lambda)y) \leq \lambda p(x) + (1-\lambda)p(y) < 1,$$

and thus $\lambda x + (1-\lambda)y \in G$.

Albeit, G is not closed. However its closure $\overline{G}$ is still convex (exercise) and absorbing (exercise).

By Theorem 20.2 there exists some $\theta > 0$ such that

$$\|x\| < \theta \quad \Rightarrow \quad x \in \overline{G}.$$

Fix now $\varepsilon > 0$. We claim that

$$\|x\| < \frac{\varepsilon}{2}\theta \quad \Rightarrow \quad p(x) < \varepsilon.$$

which readily implies continuity of p at 0.

All we need to show for this is (by homogeneity of p),

$$\|x\| < \theta \quad \Rightarrow \quad p(x) < 2. \tag{20.1}$$

Fix any such x with $\|x\| < \theta$. We know that $x \in \overline{G}$ but this does not yet imply $x \in G$ (and thus we do not know $p(x) < 1$)!.

However, since $x \in \overline{G}$ we do know that there exists some $x_1 \in G$ such that $\|x - x_1\|_X < 2^{-1}\theta$. Then

$$2(x - x_1) \in \overline{G},$$

and thus there exists x_2 with $2x_2 \in G$, i.e. $p(x_2) < \frac{1}{2}$ and

$$\|2(x - x_1 - 2x_2)\|_X < 2^{-2}\theta$$

continuing like this we find $x_n \in X$ such that $2^{n-1}x_n \in G$, i.e. $p(x_n) < 2^{1-n}$, and

$$\|x - \sum_{k=1}^{n} x_k\|_X < 2^{-n}\theta \xrightarrow{n\to\infty} 0.$$

That is $\sum_{k=1}^{\infty} x_k = x$. Then we have by countable subadditivity of p

$$p(x) \leq \sum_{k=1}^{\infty} p(x_k) < \sum_{k=1}^{\infty} 2^{1-k} = 2.$$

This implies (20.1) and we can conclude. □

Completeness of the space X is really important, Zabreiko's Lemma fails without it.

Exercise 20.10. *Consider X the vector space of sequences $(a_k)_{k\in\mathbb{N}}$ with only finitely many nonzero entries, i.e.*

$$X = \{(a_k)_{k\in\mathbb{N}} : \quad \exists K \in \mathbb{N} : a_k = 0 \quad \forall k \geq K\}.$$

We can equip X with two norms, the ℓ^∞-norm and the ℓ^1-norm.

(1) Show that $(X, \|\cdot\|_{\ell^\infty})$ *and* $(X, \|\cdot\|_{\ell^1})$ *are not complete.*
(2) Show that the
$$p((a_k)_{k\in\mathbb{N}}) := \|a_k\|_{\ell^1}$$
is a countably subadditive seminorm on $(X, \|\cdot\|_{\ell^\infty})$ *but* p *is not continuous on* $(X, \|\cdot\|_{\ell^\infty})$*.*

Zabreiko's Lemma implies several important theorems at once – we will discuss them in the next sections.

20.3. **Uniform Boundedness Principle/ Banach Steinhaus Theorem.** We have proven the *Uniform Boundedness Principle* or *Banach Steinhaus Theorem* in Theorem 12.18, but we state it here again.

Theorem 20.11 (Banach-Steinhaus or Uniform Boundedness Principle)**.** *Let* X *be a Banach space and* Y *a normed vector space. Suppose that* $\mathcal{F}$ *is a (possibly uncountable) family of continuous, linear operators* $T \in L(X,Y)$*. If* $\mathcal{F}$ *is pointwise bounded, that is if for all* $x \in X$ *we have*
$$\sup_{T\in\mathcal{F}} \|Tx\| < \infty$$
then $\mathcal{F}$ *is uniformly bounded in norm, i.e.*
$$\sup_{T\in\mathcal{F}} \|T\|$$

Exercise 20.12. *Use Zabreiko's Lemma, Lemma 20.8, to show Theorem 20.11.*

Hint: Set $p(x) = \sup_{T\in\mathcal{F}} \|Tx\|$.

Again, completeness is important

Exercise 20.13. *Let* X *be as in Exercise 20.10. For* $(a_n)_{n\in\mathbb{N}} \in X$ *and* $m \in \mathbb{N}$
$$T_m((a_n)_{n\in\mathbb{N}}) := m a_m.$$
Set
$$\mathcal{F} := \{T_m, m \in \mathbb{N}\}.$$
Show that the uniform boundedness principle fails for $\mathcal{F}$ *in* $(X, \|\cdot\|_{\ell^1})$*.*

Corollary 20.14. *Let* $(T_n)_{n\in\mathbb{N}}$ *be a sequence of linear bounded operators from a Banach space* X *to a normed space* Y*. Assume that for each fixed* $x \in X$ *the limit* $\lim_{n\to\infty} T_n x$ *exists in* Y *and set* $Tx := \lim_{n\to\infty} T_n x$*. Then* T *is a linear bounded operator from* X *to* Y*.*

Exercise 20.15. *Prove Corollary 20.14.*

20.4. **Open Mapping Theorem.** A (not necessarily linear) function from a topological X to another topological space Y is *open mapping* if for any open set $U \subset X$ we have $f(U) \subset Y$ is also open – i.e. if f maps open sets to open sets.

Theorem 20.16 (The Open Mapping Theorem)**.** *Let X, Y be two Banach spaces and $T : X \to Y$ be a bounded linear operator. If T is surjective then T is open.*

Proof. It suffices to prove that for the open unit ball $B \subset X$ the set $TB \subset Y$ is open. Indeed let $U \subset X$ be any other open set and $y \in T(U)$. Then there exists $x \in U$ such that $Tx = y$. Since $x \in U$ and U is open there exists a scaled copy of B centered at x inside U, that is there exist $r > 0$ such that

$$x + rB \subset U.$$

Since $T(B)$ is open, so is $T(x+rB) = y+rTB$. On the other hand $y+rTB = T(x+rB) \subset T(U)$ so an open neighborhood of y lies inside $T(U)$. This holds for any $y \in T(U)$ so $T(U)$ is open.

So we only show that $T(B)$ is open, where B is the open unit ball.

For $y \in Y$ set

$$p(y) := \inf\{\|x\| : \quad x \in X, Tx = y\}.$$

p is well-defined since T is surjective.

We observe that

$$p(Tx) \leq \|x\| \tag{20.2}$$

Clearly $p(0) = 0$. Let $\lambda \in \mathbb{R} \setminus \{0\}$ then

$$\begin{aligned}
p(\lambda y) =& \inf\{\|x\| : \quad x \in X, Tx = \lambda y\} = \inf\{|\lambda|\|x/\lambda\| : \quad x \in X, T\frac{1}{\lambda}x = y\} \\
=& |\lambda| \inf\{\|\tilde{x}\| : \quad \tilde{x} \in X, T\tilde{x} = y\} \\
=& |\lambda| p(y).
\end{aligned}$$

We prove countable subadditivity and triangular inequality for p at the same time.

Let $(y_n)_{n\in\mathbb{N}}$ be a sequence in Y such that $\sum_n y_n$ converges in Y. Fix $\varepsilon > 0$, then for each $n \in \mathbb{N}$ there exists $x_n \in X$ such that $Tx_n = y_n$ and

$$\|x_n\| \leq p(y_n) + 2^{-n}\varepsilon.$$

Then we have

$$\sum_n \|x_n\| \leq \sum_n p(y_n) + \varepsilon.$$

In particular $\sum_{n\in\mathbb{N}} x_n \in X$ converges (here we use that X is a Banach space). Since T is continuous we have

$$\sum_n y_n = \sum_n T(x_n) = T\left(\sum_n x_n\right)$$

Applying p we have

$$p(\sum_n y_n) = p(T\left(\sum_n x_n\right)) \overset{(20.2)}{\leq} \| \left(\sum_n x_n\right) \| \leq \sum_n \|x_n\|_X \leq \sum_n p(y_n) + \varepsilon.$$

Letting $\varepsilon \to 0$ we conclude.

By Zabreiko's Lemma, Lemma 20.8 we conclude that p is continuous. But now observe

$$T(B) = \{y \in Y : \quad \exists x \in B : \quad Tx = y\} = \{y \in Y : \quad p(y) < 1\} = p^{-1}((-1,1)).$$

Since p is continuious p^{-1} of $(-1,1)$ is open, so $T(B)$ is open, so we can conclude. □

Exercise 20.17. *Consider the identity operator on X from Exercise 20.10 defined as*

$$I\left((a_n)_{n\in\mathbb{N}}\right) := \left((a_n)_{n\in\mathbb{N}}\right).$$

Show that $I : (X, \|\cdot\|_{\ell^1}) \to (X, \|\cdot\|_{\ell^\infty})$ is a bounded linear operator that is not open.

Corollary 20.18. *Let X, Y be two Banach spaces and let $T : X \to Y$ be a compact linear operator. If T is surjective, then Y must be finite dimensional.*

Proof. If $T : X \to Y$ is surjective (it is also continuous) then T is open by Theorem 20.16.

Let B denote the open unity ball in X. Since T is open, then $T(B) \subset Y$ is an open set containing the origin, so there exists a closed ball $\overline{B(0,\varepsilon)} \subset T(B)$. Now $\overline{B(0,\varepsilon)}$ is compact (since its closed and lies in the impage of a bounded set under the compact operator T). The Riesz lemma, Exercise 7.20, says that the unit ball (hence also $\overline{B(0,\varepsilon)}$) is only compact iff the dimension is finite. □

20.5. **Inverse Mapping Theorem.** As a corollary of the open mapping theorem we have the the inverse mapping theorem. Observe that in general just because $f : X \to Y$ is continuous and bijective there is (in general) no reason that f^{-1} is continuous, Exercise 20.20.

For linear maps on Banach spaces this is true however.

Theorem 20.19 (Bounded inverse theorem / Inverse Mapping theorem)**.** *Let X be a Banach space and $T : X \to Y$ be a bounded linear operator such that $T : X \to Y$ is bijective. Then $T^{-1} : Y \to X$ is a linear bounded map*

Proof. The fact that T^{-1} is linear is easy to show from the bijectivity. By the open mapping theorem T maps open sets to open sets, and since $(T^{-1})^{-1} = T$ we have that T^{-1} is necessarily continuous. □

Exercise 20.20. *Consider the identity operator on X from Exercise 20.10 defined as*

$$I\left((a_n)_{n\in\mathbb{N}}\right) := \left((a_n)_{n\in\mathbb{N}}\right).$$

Show that $I : (X, \|\cdot\|_{\ell^1}) \to (X, \|\cdot\|_{\ell^\infty})$ is a bounded linear operator which is bijective – but whose inverse is not continuous.

The inverse mapping theorem says that vectors spaces X which are Banach with respect to two norms which create the same topology have equivalent norms.

Exercise 20.21. *Let $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ be two Banach spaces with the same underlying Vector space X. If the identity map $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is continuous, then the norms $\|\cdot\|_1$ is equivalent to $\|\cdot\|_2$.*

Exercise 20.22. *Let $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ be two Banach spaces with the same underlying Vector space X. Assume that for some $\Lambda > 0$ we have*

$$\|x\|_1 \leq \Lambda \|x\|_2 \quad \forall x \in X.$$

Then there exists $\lambda > 0$ such that

$$\lambda \|x\|_2 \leq \|x\|_1 \quad \forall x \in X.$$

20.6. Closed Graph theorem.

Theorem 20.23 (Closed Graph Theorem)**.** *Let $T : X \to Y$ be a linear map from a Banach space X to a Banach space Y.*

Suppose that T is a closed operator in the following sense: whenever $x_n \in X$ converges to $x \in X$ and at the same time Tx_n converges to $y \in Y$ then we have $Tx = y$.

Then T is a bounded operator.

Proof. For $x \in X$ we set $p(x) := \|Tx\|_Y$. If we show that p is continuous then there must be some $\delta > 0$ such that

$$\sup_{\|x\| \leq \delta} \|Tx\|_Y \equiv \sup_{\|x\|_X \leq \delta} (p(x) - p(0)) \leq 1.$$

This clearly implies

$$\|Tx\|_Y \leq \frac{1}{\delta} \|x\|_X,$$

and thus T is bounded.

How do we show p is continuous? Zabreiko's lemma, Lemma 20.8.

Clearly $p(\lambda x) = |\lambda| p(x)$.

So let $\sum_{k \in \mathbb{N}} x_k$ be a convergent series in X. We need to show

$$p(\sum_k x_k) \leq \sum_k p(x_k).$$

If the right-hand side is infinite, there is nothing to show. So we may assume that

$$\sum_k p(x_k) = \sum_k \|Tx_k\|_Y < \infty.$$

Since Y is a Banach space we have that $\sum_k Tx_k$ converges in Y. Then we have the following

$$\sum_{k=1}^{n} x_k \xrightarrow{n\to\infty} \sum_{k=1}^{\infty} x_k \quad \text{in } X$$

and

$$T\left(\sum_{k=1}^{n} x_k\right) = \sum_{k=1}^{n} Tx_k \xrightarrow{n\to\infty} \sum_{k=1}^{\infty} Tx_k \quad \text{in } Y$$

By the assumption of closedness of T we then have

$$T(\sum_{k=1}^{\infty} x_k) = \sum_{k=1}^{\infty} Tx_k,$$

and thus

$$p(\sum_{k=1}^{\infty} x_k) = \|T(\sum_{k=1}^{\infty} x_k)\|_Y = \|\sum_{k=1}^{\infty} Tx_k\|_Y \leq \sum_{k=1}^{\infty} \|Tx_k\|_Y = \sum_{k=1}^{\infty} p(x_k).$$

□

We can strengthen the Inverse mapping theorem Theorem 20.19,

Theorem 20.24. *Let X, Y be Banach spaces, and assume $Z \subset X$. Assume that $T : Z \to Y$ is linear, closed (as defined in Theorem 20.23), injective, and surjective. Then there exists a map $T^{-1} \in \mathcal{L}(Y, X)$ such that $TT^{-1} : Y \to Y$ is the identity and $T^{-1}T : Z \to Z$ is the identity.*

A good example to have in mind for the above is $X = L^2(\Omega)$, $Z = W^{2,2}(\Omega) \cap W_0^{1,2}(\Omega)$ (equipped with the L^2-norm!) and $Y = L^2(\mathbb{R}^d)$ and $T = \Delta$.

Proof of Theorem 20.24. Set

$$\Gamma := \{(x, Tx) : \quad x \in Z\} \subset X \times Y.$$

Since T is a closed operator we have that $\Gamma \subset X \times Y$ is a closed subspace, and thus a Banach space. Set $\pi : \Gamma \to Y$,

$$\pi((x, y)) := y$$

the projection into Y. Since T is injective, so is π (on Γ). Since T is surjective, so is π. That is $\pi : \Gamma \to Y$ is bijective.

Moreover $\pi : \Gamma \to Y$ is closed. If $(x_k, y_k) \in \Gamma$ converge to $(x, y) \in \Gamma$ and $z_k \in Y$ converge to $z \in Y$ and we know that $\pi((x_k, y_k)) = z_k$ then we have $y_k = z_k$ and thus $y = z$. Since $(x_k, y_k) \in \Gamma$ we have $Tx_k = y_k$, and by closedness of T we conclude that $Tx = y$. Thus $\pi(x, y) = y = z$, and π is closed.

By the closed graph theorem, Theorem 20.23, π is continuous. By Theorem 20.19 $\pi^{-1} : Y \to \Gamma$ is continuous.

Set

$$\psi : \Gamma \to X$$

as the projection $\psi(x,y) := x$. Clearly $\psi : \Gamma \to X$ is bounded linear.

Now we check that

$$T^{-1} := \psi \circ \pi^{-1} : Y \to X$$

is indeed the map we are looking for. □

Exercise 20.25. *Consider the identity operator on X from Exercise 20.10 defined as*

$$I\left((a_n)_{n\in\mathbb{N}}\right) := \left((a_n)_{n\in\mathbb{N}}\right).$$

Show that $I : (X, \|\cdot\|_{\ell^\infty}) \to (X, \|\cdot\|_{\ell^1})$ contradicts the closed graph theorem on incomplete normed spaces.

Exercise 20.26. *Show that Closed Graph theorem, Inverse Mapping theorem, Open Map theorem are equivalent (i.e. show that assuming any of the three is proven, it implies the other two).*

An application of the closed graph theorem

Proposition 20.27 (Hellinger-Toeplitz)**.** *Let H be a Hilbert space and suppose that $T : X \to X$ is a linear (but not necessarily bounded) map, such that there exists the adjoint T^*, namely a map $T^* : H \to H$ such that*

$$\langle Tx, y\rangle = \langle x, T^* y\rangle \quad \textit{for all } x, y \in H.$$

Then T is bounded.

Exercise 20.28. *Show that T^* as above is necessarily linear*

Proof. In view of Theorem 20.23 we need to show that T is closed. So assume that $(x_n)_{n\in\mathbb{N}}$ converges in H to x, and $Tx_n \xrightarrow{n\to\infty} z$. We need to show that $Tx = z$.

We have for any $y \in H$,

$$\langle z, y\rangle = \lim_{n\to\infty} \langle Tx_n, y\rangle = \lim_{n\to\infty} \langle x_n, T^* y\rangle = \langle x, T^* y\rangle = \langle Tx, y\rangle$$

That is

$$\langle z - Tx, y\rangle = 0 \quad \forall y \in H.$$

Taking $y = z - Tx$ we find

$$\|z - Tx\|_H^2 = 0,$$

thus $z = Tx$ and we can conclude. □

Exercise 20.29 (Jesús Gil de Lamadrid)**.** *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be both Banach spaces. Assume Y can be equipped with another norm $\|\cdot\|_2$ which satisfies, for some $C > 0$*

$$\|x\|_2 \le C\, \|x\|_Y \quad \forall x \in Y.$$

(1) *Give an example to show that $(Y, \|\cdot\|_2)$ may not be complete.*

(2) Let $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_2)$ is a bounded linear operator. Using the closed graph theorem, show that $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is also bounded.

Exercise 20.30. *(1) Let $(X, \|\cdot\|_X)$ be a Banach space and $Y \subset X$ a Vector-subspace of X. Assume that there exists a norm $\|\cdot\|_Y$ which satisfies*

$$\|y\|_X \leq C\,\|y\|_Y \quad \text{for all } y \in Y.$$

Let $T : (X, \|\cdot\|_X) \to (X, \|\cdot\|_X)$ be a bounded linear operator, and assume that $T(Y) \subset Y$. Then

$$T\Big|_Y : (Y, \|\cdot\|_Y) \to (Y, \|\cdot\|_Y)$$

is a bounded linear operator.

Hint: Show that $T\Big|_Y : (Y, \|\cdot\|_Y) \to (Y, \|\cdot\|_X)$ is bounded and use Exercise 20.29.

(2) Let $T : (L^2([0,1]), \|\cdot\|_{L^2([0,1])}) \to (L^2([0,1]), \|\cdot\|_{L^2([0,1])})$ be a linear bounded operator with the following property. If $\varphi \in C^0([0,1])$ then $T\varphi \in C^0([0,1])$. Then

$$T : (C^0([0,1]), \|\cdot\|_{L^\infty([0,1])}) \to (C^0([0,1]), \|\cdot\|_{L^\infty([0,1])})$$

is a linear bounded operator.

21. Closed Range Theorem, Spectral Theory, Fredholm Alternative

Although we did not discuss this before, especially for spectral theory it makes sense to consider spaces X on the compact field $\mathbb{C}$. There is not much difference (except for the scalar product) in doing so.

21.1. Adjoint operators.

Let $T : X \to Y$ be a linear and bounded operator.

Consider $y^* \in Y^*$. Then $y^* \circ T \in X^*$. We call the operator that maps y^* into $y^* \circ T$ the *adjoint* operator T^*,

$$T^* : Y^* \to X^*,$$

and we record its defining property

$$y^*[Tx] = (T^* y^*)[x] \quad \forall x \in X.$$

Since T is linear, so is T^*, and we have

Theorem 21.1. *Let $T \in \mathcal{L}(X, Y)$. Then $T^* \in L(Y^*, X^*)$ and we have*

$$\|T\|_{\mathcal{L}(X,Y)} = \|T^*\|_{\mathcal{L}(Y^*,X^*)}$$

Proof. For $y^* \in Y^*$ we have

$$|y^*[Tx]| \leq \|y^*\|_{Y^*}\,\|T\|_{\mathcal{L}(X,Y)}\,\|x\|_X.$$

Thus,

$$\begin{aligned}\sup_{\|y^*\|_{Y^*}\le 1} \|T^*y^*\|_{X^*} =& \sup_{\|x\|_X\le 1,\ \|y^*\|_{Y^*}\le 1} |T^*y^*[x]|\\ =& \sup_{\|x\|_X\le 1,\ \|y^*\|_{Y^*}\le 1} |y^*[Tx]|\\ =&\|T\|_{\mathcal{L}(X,Y)}\end{aligned}$$

by Corollary 10.14. □

How should we think of T^*?

Assume that X and Y are finitedimensional spaces, e.g. $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$. Let $T : X \to Y$ is linear (and thus bounded). Every norm is equivalent on finite dimensional spaces, so we might take as well the scalar product norm. We then have that each element of X^* corresponds to an element of X, via the identification

$$x^* : \quad X \ni z \mapsto \langle z, x\rangle$$

So if $y^* \in Y^*$ then we have $y^* = \langle\cdot, y\rangle$. Thus

$$y^*[Tx] = (T^*y^*)[x] \quad \forall x \in X, y^* \in Y^*$$

becomes

$$\langle y, Tx\rangle_{\mathbb{R}^m} = \langle T^*y, x\rangle \quad \forall x \in X, y \in Y$$

Now in finite dimensions, we know that T corresponds to a matrix $T \in R^{m\times n}$ and Tx is simply the matrix-vector product. Then what is T^*? it is the transpose $T^t \in \mathbb{R}^{m\times n}$.

$$\langle y, Tx\rangle_{\mathbb{R}^m} \equiv y^tTx = (T^ty)\,x \equiv \langle T^*y, x\rangle_{\mathbb{R}^m}.$$

So T^* is a generalization of the transpose operator in $\mathbb{R}$-valued vector spaces.

What if $X = \mathbb{C}^n$ and $Y = \mathbb{C}^m$? Recall that the complex scalar product

$$\langle v, w\rangle_{\mathbb{C}^m} = v^t\overline{w}$$

Then Then we have

$$y^t\overline{Tx} = (T^*y)^t\overline{x},$$

so $T^* = \overline{T}^t$ – i.e. the *conjugate transpose* or *Hermitian transpose*.

More generally, if $T : H_1 \to H_2$ are two Hilbert spaces, we have $H_1^* \cong H_1$, and $H_2^* \cong H_2$, see Corollary 19.19, so we can identify $T^* : H_2 \to H_1$ so that

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

Example 21.2.

Let $I : X \to X$ be the identity. Then I^* is the identity as well.

Let $X = Y = \ell^p(\mathbb{N})$ where $1 \le p < \infty$. Set the *right-shit operator*

$$S_+ : (x_1, x_2, x_3, \ldots) \mapsto (0, x_1, x_2, x_3, \ldots).$$

What is S_+^*? By Riesz representation theorem we know that any element of $(\ell^p)^*$ can be identified with an element of ℓ^q, where $\frac{1}{p} + \frac{1}{q} = 1$.

Given $x \in \ell^p(\mathbb{N})$, $y \in \ell^q(\mathbb{N})$, we have

$$\sum_{k=1}^{\infty} \left((S_+)x\right)_k y_k = \sum_{k=2}^{\infty} x_{k-1} y_k = \sum_{\tilde{k}=1}^{\infty} x_{\tilde{k}} y_{\tilde{k}+1} = \sum_{k=1}^{\infty} x_k (S_+)^* y_k.$$

That is $(S_+)^*$ is the left-shit operator

$$S_- : (x_1, x_2, x_3, \ldots) \mapsto (x_2, x_3, \ldots).$$

Exercise 21.3. *Let X, Y, Z be normed vector spaces, $T : X \to Y$, $S : Y \to Z$ be linear and bounded. Show that*

(1) $(T_1 + T_2)^* = T_1^* + T_2^*$
(2) $(\lambda T_1)^* = \lambda T_1^*$ *for all* $\lambda \in \mathbb{C}$,
(3) $(S \circ T)^* = T^* \circ S^*$

Theorem 21.4. *Let $T : X \to Y$ be linear and bounded, and assume $T^{-1} : Y \to X$ exists and is continuous. Then T^* is continously invertible as well and we have*

$$(T^*)^{-1} = (T^{-1})^*.$$

Proof. Since T is continuously invertable we have $T \circ T^{-1} = Id_X$ and $T^{-1}T = Id_Y$. We take the adjoint of both sides, and from Exercise 21.3 we obtain

$$T^* \circ \left(T^{-1}\right)^* = I_{X^*}$$

$$\left(T^{-1}\right)^* \circ T^* = I_{Y^*}$$

Thus $(T^{-1})^*$ is the inverse of T^*. $\square$

21.2. **Closed Range Theorem.** We can use the adjoint operator to study solutions of the equation

$$Tx = y$$

Solution means that for given y we find x and hopefully x is unique – i.e. the question is about surjectivity and injectivity. Recall the finite-dimensional case: If $T : V \to W$ is a matrix operator then T is surjective if and only if $T^t \equiv T^*$ is injective – this is the rank-nullity theorem:

$$\dim \mathrm{im} T + \dim \ker T = \dim V.$$

For an operator $T : X \to Y$ we denote by its *image* or *range*

$$\mathrm{im} T = \{y \in Y : \quad \exists x \in X : Tx = y\}.$$

and its *kernel*

$$\ker T = \{x \in X : \quad Tx = 0\}.$$

Observe that the $\ker T$ is always a closed set if T is continuous – but $\mathrm{im} T$ may not be closed.

The following result is an extension of that concept – where we will use of *annihilators*: For $A^* \subset Y^*$

$$^\perp A^* = \{y \in Y : \text{for all } y^* \in A^*\text{: } y^*[y] = 0\}$$

and for $A \subset Y$

$$A^\perp := \{y^* \in Y^* : \forall y \in Y : y^*[y] = 0\}.$$

We then have (cf. Lemma 19.9 for the Hilbert-space version)

Exercise 21.5. *(1) Let $A \subset X$. Then $A^\perp \subset X^*$ is a closed linear subspace of X^**
(2) Let $A \subset X^*$ then $^\perp A \subset X$ is a closed linear subspace of X

Observe that for Hilbert spaces H its dual H^* corresponds to the scalar product with H itself.

Exercise 21.6. *Let H be a Hilbert space and consider the canonical identification $\iota : H^* \to H$ from the Riesz representation theorem, Theorem 19.18. Let $A \subset H$, and consider by the usual orthogonal space $\tilde{A}$*

$$\tilde{A} := \{x \in H : x \perp a \quad \forall a \in \tilde{A}\}$$

Show that

$$\tilde{A} = \iota A^\perp$$

We also have the following (cf. Lemma 19.14 for the Hilbert-space version).

Exercise 21.7. *Let X be a normed vector space and $U \subset X$ be a subspace. Then*

$$^\perp\left(U^\perp\right) = \overline{U}$$

Hint: If $x \notin \overline{U}$ use the Hahn-Banach theorem, Corollary 10.18. For the other direction show that $^\perp\left(U^\perp\right)$ is closed and $U \subset {}^\perp\left(U^\perp\right)$.

With the notion of annihilators we have the following partial extension of the rank-nullity theorem:

Theorem 21.8. *Let X and Y be normed vector spaces and $T : X \to Y$ be linear and bounded. Then*

(1) $(\mathrm{im} T)^\perp = \ker T^*$
(2) $^\perp(\ker T^*) = \overline{\mathrm{im} T}$
(3) $^\perp(\mathrm{im} T^*) = \ker T$
(4) $(\ker T)^\perp \supset \overline{\mathrm{im} T^*}$

Proof of Theorem 21.8(1). Let $y^* \in (\mathrm{im}T)^{\perp}$. Then we have for all $x \in X$

$$0 = y^*[Tx] = (T^* y^*)(x).$$

Since this holds for all x we have that $T^* y^* = 0$ and thus $y^* \in \ker T^*$.

On the other hand, assume that $y^* \in \ker T^*$. Then $y^*[Tx] = 0$ for all $x \in X$ and thus $y^* \in (\mathrm{im}T)^{\perp}$. □

Proof of Theorem 21.8(2). From (1) we have

$$(\mathrm{im}T)^{\perp} = \ker T^*$$

Applying $^{\perp}$ to both sides we have

$$^{\perp}\left((\mathrm{im}T)^{\perp}\right) = {}^{\perp}(\ker T^*)$$

So the claim follows once we show that

$$^{\perp}\left((\mathrm{im}T)^{\perp}\right) = \overline{\mathrm{im}T}.$$

This follows from Exercise 21.7 □

Proof of Theorem 21.8(3),(4). We skip those, they are similar to (1) and (2) – cf. [Clason, 2020, Theorem 9.6] □

In particular, T is surjective if and only if T^* is injective and $\mathrm{im}T$ is closed.

Theorem 21.9 (Closed range theorem)**.** *Let X, Y be Banach spaces and assume that $T : X \to Y$ is a linear bounded operator. Then the following properties are equivalent.*

(1) $\mathrm{im}T \subset Y$ *is closed*
(2) $\mathrm{im}T = {}^{\perp}\ker T^*$
(3) $\ker(T^*)$ *is closed*
(4) $\mathrm{im}T^* = (\ker T)^{\perp}$.

Proof of Theorem 21.9: (2) ⇒ (1). By Exercise 21.5, if (2) holds then $\mathrm{im}T$ is closed. □

Proof of Theorem 21.9: (1) ⇒ (2). Let $y \in \mathrm{im}T$. Then there exists $x \in X$ such that $Tx = y$. For any $z^* \in \ker(T^*) \subset Y^*$ we then find

$$z^*[y] = z^*[Tx] = (T^* z^*)[x] \overset{z^* \in \ker T^*}{=} 0.$$

Thus $y \in {}^{\perp}\ker T^*$.

So we have shown $\mathrm{im}T \subset {}^{\perp}\ker T^*$.

For the other direction, it follows from Exercise 21.7, but we repeat the proof: set $Z := {}^{\perp}\ker T^* \subset Y^*$ (a linear normed space) and assume there exists $y_0 \in Z \setminus \mathrm{im}T$. Then

$$y^*[y_0] = 0 \quad \forall y^* \in \ker T^*.$$

On the other hand we can apply Hahn-Banach theorem, in the form of Corollary 10.18 (here we assume (1) i.e. that $\mathrm{im}T \subset Y$ is a closed space). Then we find some $y^* \in Z^*$ with $y^*[y_0] \neq 0$ and $y^*(\mathrm{im}T) = 0$. The latter says

$$y^*[Tx] = 0 \quad \forall x \in X.$$

We can extend (again by Hahn-Banach) y^* from $\mathrm{im}T$ to a map on Y, i.e. $y^* \in Y^*$. Thus,

$$(T^* y^*)[x] = y^*[Tx] = 0 \quad \forall x \in X.$$

But this implies that $T^* y^* = 0$, thus $y^* \in \ker T^*$. Since $y_0 \in Z = {}^\perp \ker T^*$ we have

$$y^*[y_0] = 0,$$

so we found a contradiction and such a y_0 does not exist.

We can conclude. □

Proof. The proof of the remaining directions of Theorem 21.9 can be found e.g. in [Clason, 2020, Theorem 9.10]. □

As a corollary to Theorem 21.9 we obtain

Corollary 21.10. *Let X, Y be Banach spaces and assume that $T : X \to Y$ is a bounded linear operator with dual operator $T^* : Y^* \to X^*$. Assume that*

(1) $\mathrm{im}T \subset Y$
(2) $\ker T^* = 0$.

Then T is surjective. I.e. for any $y \in Y$ there exists $x \in X$ such that $Tx = y$.

If T is moreover injective, we have for some constant $\mu > 0$ (only depending on T)

$$\|x\| \leq \mu \|y\|.$$

Proof. The surjectivity follows immediately from Theorem 21.9. The last estimate follows from Theorem 20.24 (Inverse mapping theorem). □

Remark 21.11. The closed range theorem can extended to densely defined operators (such as Δ on L^2) – this can be be found in [Brezis, 2011, Theorem 2.19.]

21.3. **Example: Solving a PDE via the closed range theorem.** The typical example to have in mind is the following.

Example 21.12. Let $\Omega \subset\subset \mathbb{R}^n$ and $A : \Omega \to \mathbb{R}^{n\times n}$ an measurable, bounded, and uniformly elliptic matrix i.e. for some $\lambda > 0$

$$\langle Av, v\rangle \geq \lambda |v|^2 \quad \text{a.e. in } \Omega$$

and A is bounded $\sup_\Omega |A| \leq \Lambda$.

We want to solve

$$\begin{cases} -\mathrm{div}\,(A\nabla u) = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

We already discussed how to obtain a solution $u \in W^{1,2}_0(\Omega)$ via the Riesz representation theorem (in the form of Lax-Milgram Example 19.25).

Recall that

$$H^{-1}(\Omega) := \left(W^{1,2}_0(\Omega)\right)^*.$$

This is a Hilbert-space, so $(H^{-1}(\Omega))^*$ "$=$" $H^{-1}(\Omega)$ – but this is inconvenient to use. Instead we use that $W^{1,2}_0(\Omega)$ is reflexive (as a Hilbert space), so we use $(H^{-1}(\Omega))^* = W^{1,2}_0(\Omega)$.

Set

$$Tu := \mathrm{div}\,(A\nabla u).$$

Then $T : W^{1,2}(\Omega) \to H^{-1}(\Omega)$ is *continuous*.

Let $v \in (H^{-1}(\Omega))^*$ then we can identify v with an element in $W^{1,2}_0(\Omega)$, and we have

$$v[Tu] = \int_\Omega \langle A\nabla u, \nabla v\rangle = \int_\Omega \langle \nabla u, A^t \nabla v\rangle = -\mathrm{div}\,(A^t\nabla v)[u].$$

That is $T^* : W^{1,2}_0(\Omega) = (H^{-1}(\Omega))^* \to H^{-1}(\Omega) = (W^{1,2}(\Omega))^*$ is simply $-\mathrm{div}\,(A^t\nabla\cdot)$.

- It is easy to see that both T and T^* are injective, i.e. $\ker T^* = 0$ (this follows from uniqueness of solutions).
- We show that $T(W^{1,2}_0(\Omega))$ is a closed subset of $H^{-1}(\Omega)$: Assume $\Delta u_k = f_k$ for $u_k \in W^{1,2}_0$ and $f_k \in H^{-1}$ and $f_k \to f$ in H^{-1}. Testing this equation with u_k we have
$$\|\nabla u_k\|^2_{L^2(\Omega)} \leq \|f_k\|_{H^{-1}}\, \|u_k\|_{W^{1,2}}.$$
Using Poincaré inequality we obtain
$$\|u_k\|^2_{W^{1,2}(\Omega)} \lesssim \|f_k\|_{H^{-1}}\, \|u_k\|_{W^{1,2}}.$$
That is
$$\|u_k\|_{W^{1,2}(\Omega)} \lesssim \|f_k\|_{H^{-1}}.$$
That is u_k is bounded in $W^{1,2}$, so up to subsequence u_k weakly converges to some $u \in W^{1,2}_0(\Omega)$. Using testfunction we see that $\Delta u = f$. Thus, $\Delta(u_k - u) = f_k - f$, so from the same argument as above
$$\|u_k - u\|_{W^{1,2}(\Omega)} \lesssim \|f_k - f\|_{H^{-1}} \xrightarrow{k\to\infty} 0.$$

Corollary 21.10 implies surjectivity of T, i.e. for any $f \in H^{-1}(\Omega)$ there exists $u \in W^{1,2}_0(\Omega)$ solving

$$\mathrm{div}\,(A\nabla u) = f.$$

21.4. **Fredholm Alternative.** The following looks very similar to the closed range theorem, Theorem 21.9, (and philosophically it is not too far). Our proof will however not use that.

Theorem 21.13 (Fredholm Alternative)**.** *Let X be a Banach space. Let $K : X \to X$ be a compact operator and $\lambda \in \mathbb{C} \setminus \{0\}$. Then one and only one of the following properties holds.*

(1) The homogeneous equation
$$\lambda x - Kx = 0$$
has a nontrivial solution $x \neq 0$
(2) For each $y \in X$ the inhomogeneous
$$\lambda x - Kx = y$$
has a unique solution $x \in X$

Proof of Theorem 21.13: if (1) holds then (2) does not hold. This is the easy part: Assume (1) holds, that is there exists $\bar{x} \neq 0$ such that
$$\lambda \bar{x} - K\bar{x} = 0.$$
Take any $y \in X$. Either there is *no* solution to
$$\lambda x - Kx = y$$
(i.e. *existence fails*). Alternatively, if there is some $\tilde{x}$ with
$$\lambda \tilde{x} - K\tilde{x} = y$$
then also $\tilde{x} + \bar{x}$ solves the same equation (i.e. *uniqueness fails*). In either case (2) cannot be true. □

For the other direction in the proof of Theorem 21.13 we need two lemmata:

Lemma 21.14. *Let $K : X \to X$ linear bounded and compact and set $Y := \ker(I - K)$. For each $x \in X$ there exists $y \in Y$ such that*
$$\|x - y\| \leq \Lambda \|(I - K)x\|.$$

Proof. First we observe that for each $x \in X$ there exists (at least) one $y \in Y$ such that
$$\|y - x\| = \inf_{y \in Y} \|x - y\|.$$
Indeed, let $(y_\ell)_{\ell \in \mathbb{N}}$ be a sequence
$$\|y_\ell - x\| \xrightarrow{\ell \to \infty} \inf_{y \in Y} \|x - y\|.$$
Then w.l.o.g.
$$\|y_\ell\| \leq \|x\| + \inf_{y \in Y} \|x - y\| + 1 \quad \forall \ell$$

That is y_ℓ is bounded. Moreover $y_\ell = Ky_\ell$ (since $y_\ell \in Y$), so by compactness, up to subsequence, $y_\ell \xrightarrow{\ell\to\infty} y \in \overline{Y}$. Since Y is closed (by continuity of I and K) we find that $y \in Y$, and indeed

$$\|y - x\| = \inf_{y\in Y} \|x - y\|.$$

Assume now the claimed inequality does not hold for any Λ, then there exists x_k such that

$$\inf_{y\in Y} \|x_k - y\| > k\|(I-K)x_k\|.$$

Taking y_k the optimal projection of x_k into Y from above (i.e. $\|y_x - x_k\| = \inf_{y\in Y}\|y - x\|$) we can also say

$$\|(x_k - y_k) - y\| \geq \|x_k - y_k\| > k\|(I-K)(x_k - y_k)\| \quad \forall y \in Y.$$

Set $z_k := \frac{x_k - y_k}{\|x_k - y_k\|}$, then we conclude

$$\|z_k - y\| \geq 1 > k\|(I-K)z_k\| \quad \forall y \in Y.$$

Since K is compact, we can pick a subsequence such that $Kz_k \xrightarrow{k\to\infty} \bar{z}$, then from the above inequality we find

$$\|z_k - \bar{z}\| \leq \|(I-K)z_k\| + \|Kz_k - \bar{z}\| \xrightarrow{k\to\infty} 0.$$

That is $\lim_{k\to\infty} z_k = \bar{z}$, and thus

$$\|\bar{z} - y\| \geq 1 = \|\bar{z}\| \quad \forall y \in Y.$$

On the other hand from the convergence we have $\bar{z} = \lim_{k\to\infty} Kz_k = K\bar{z}$, that is $\bar{z} \in Y$. This contradicts the previous estimate since we could choose $y = \bar{z}$. $\square$

Lemma 21.15. *Let X be a Banach space and $K : X \to X$ compact and bounded and linear. Then* $\operatorname{im}(I-K)$ *is closed.*

Proof. Let $z_k \in \operatorname{im}(I-K)$ with $z = \lim_{k\to\infty} z_k$. We need to find $x \in X$ with $(I-K)x = z$. Firstly, there must be $x_k \in X$ such that $z_k = (I-K)x_k$. Secondly, by Lemma 21.14 we may assume that

$$\|x_k\| \leq \Lambda_K \|z_k\|,$$

that is $(x_k)_{k\in\mathbb{N}}$ is bounded. Up to subsequence we then may assume that Kx_k converges, and thus $x_k = z_k + Kx_k$ is convergent to some $x \in X$. By continuity we have $(I-K)x = z$. $\square$

Proof of Theorem 21.13: if (1) does not hold then (2) holds. W.l.o.g. $\lambda = 1$, otherwise multiply with $\frac{1}{\lambda}$ and consider $\frac{1}{\lambda}K$ instead (which is still compact).

We have $S := I - K : X \to X$ is a linear bounded operator, which by assumption is injective. We need to show that S is surjective.

Set

$$S^n := S \circ \ldots \circ S,$$

and set $U^n := \mathrm{im} S^n$. By Lemma 21.15, U^n are closed subspaces, since we can write

$$S^n = (I-K)^n = I + \sum_{\ell=0}^{n} \binom{n}{\ell} (-K)^\ell = I + \tilde{K},$$

where $\tilde{K}$ is compact.

Clearly $U_{n+1} = S^{n+1}(X) = S^n(S(X)) \subset S^n(X) = U_n$. If S is not surjective, then this set-inequality is strict:

Indeed assume that $y \not\in \mathrm{im} S$. Then $S^n y \in U_n$. Assume $S^n y \in U_{n+1}$ then there must be some $x \in X$ with $S^{n+1}x = S^n y$, which implies

$$S^n(Sx - y) = 0.$$

Since S is injective by assumption, we conclude that $Sx = y$, a contradiction to $y \not\in \mathrm{im} S$.

So either S is surjective, or there exists $x \not\in \mathrm{im} S$, which implies that $U_{n+1} \subsetneq U_n$.

Hence, if we assume S is not injective, we can apply the Riesz Lemma, Lemma 7.18, to the closed subspaces $U_{n+1} \subsetneq U_n$: we can pick sequences $(x_n)_{n\in\mathbb{N}}$ such that $x_n \in U_n$, $\|x_n\|_X = 1$

$$\|u - x_n\|_X \geq \frac{1}{2} \quad \forall u \in U_{n+1}.$$

Thus if $m > n$

$$\|Kx_n - Kx_m\|_X = \| \underbrace{Sx_n - Sx_m + x_m}_{\in U_{n+1}} - x_n\| \geq \frac{1}{2}.$$

But this implies that K cannot be compact, because the sequence is $(x_n)_{n\in\mathbb{N}}$ are bounded. Contradiction, so S must have been surjective all along. □

Theorem 21.16. *Let $\Omega \subset \mathbb{R}^n$ be a smoothly bounded open set. Let $\lambda \in L^\infty(\Omega)$*

Then exactly one of the following statements is true:

- $$\Delta u - \lambda(x) u = 0 \quad \text{in } \Omega$$

  has a nontrivial solution $u \in H^1_0(\Omega) \setminus \{0\}$ in distributional sense (i.e. λ behaves like an eigenvalue!)
- *For each $f \in L^2(\Omega)$ there exists exactly on $u \in H^1_0(\Omega)$ such that*

  $$\Delta u - \lambda(x) u = f \quad \text{in } \Omega$$

  in the distributional sense.

Proof. Fix $\lambda \in L^\infty(\Omega)$.

We want to apply the Fredholm alternative, but where is the compact operator?

And why don't we just apply the Direct method, Theorem 12.30, to solve this sort of PDE? Well the issue is that λ has no prescribed sign, so the energy we used in Theorem 12.30 may not be coercive.

Let $\Lambda := \|\lambda\|_{L^\infty(\Omega)}$

$$\mathcal{E}(v) := \frac{1}{2}\int_\Omega |Dv|^2 + \frac{1}{2}\int_\Omega (\lambda(x)+\Lambda)\,|v|^2 + \int fv$$

This energy is coercive, and as in Theorem 12.30, for each $f \in L^2(\Omega)$ there exists a unique solution $u = u(f) \in H_0^1(\Omega)$ to its Euler-Lagrange equation

$$\Delta u - (\lambda + \Lambda)\,u = f \quad \text{in } \Omega,$$

and we have

$$\|\nabla u\|_{L^2(\Omega)} \leq C\,\|f\|_{L^2(\Omega)}$$

So if we set $Kf := u(f)$ we have a linear bounded map from $L^2(\Omega) \to H_0^1(\Omega)$. By Rellich's theorem, Theorem 13.35,we can consider $K : L^2(\Omega) \to L^2(\Omega)$ as a linear bounded, *compact*, map.

Our original solution

$$\Delta u - \lambda u = f$$

is then equivalent to

$$\Delta u - (\lambda + \Lambda)u = f - \Lambda u$$

Which, applying K to this equation, is equivalent to

$$u = K(f - \Lambda u)$$

or equivalently,

$$u + \Lambda K u = Kf.$$

Setting $h := Kf$ we conclude that we would like to solve

$$u + \Lambda K u = h.$$

The claim now follows from Theorem 21.13. □

21.5. **Spectrum.** (We follow substantially the exposition in [Clason, 2020]) For a matrix $A \in \mathbb{C}^{n\times n}$ we say that $\lambda \in \mathbb{C}$ is an eigenvalue if there exists an *eigenvector* $v \in \mathbb{C}^n \setminus \{0\}$

$$Av = \lambda v$$

We now extend this notion to bounded linear operators (where we switch for completeness now to $\mathbb{C}$ as an underlying field).

Let $T \in \mathcal{L}(X, X)$. Then the *spectrum* $\sigma(T)$ is defined as

$$\sigma(T) = \{\lambda \in \mathbb{C} : \quad T - \lambda I \quad \text{is not bijective}\},$$

that is $\lambda \not\in \sigma(T)$ if and only if $T - \lambda I : X \to X$ is bijective (and thus by the inverse mapping theorem, Theorem 20.19, $(T - \lambda I)^{-1} : X \to X$ is a linear bounded operator).

If for some $\lambda \in \mathbb{C}$ there exists $v \in X \setminus \{0\}$ such that

$$Tv = \lambda v,$$

then we call λ an *eigenvalue* and v its *eigenvector*. The *Eigenspace* of an eigenvalue λ is given by

$$\ker(T - \lambda I) := \{v \in X : Tv - \lambda v = 0\}.$$

In contrast to the matrix (i.e. finite dimensional) case, $\lambda \in \sigma(T)$ does *not* necessarily mean λ is an eigenvalue! It could also happen that $T - \lambda I$ is simply not surjective!

The complement of the spectrum (i.e. all λ where $T - \lambda I$ is continuously invertible is called the *resolvent set*

$$\rho(T) = \mathbb{C} \setminus \sigma(T).$$

and

$$R(\lambda) = (T - \lambda I)^{-1}$$

is the *resolvent function*.

The spectrum $\sigma(T)$ is sometimes decomposed into

(1) *point spectrum* $\sigma_p(T) = \{\lambda \in \mathbb{C} : \lambda I - T \quad \text{not injective}\}$
(2) *continuous spectrum* $\sigma_c(T) := \{\lambda \in \mathbb{C} : \lambda I - T \quad$ is injective but not surjective, but with dense range
(3) *residual spectrum* $\sigma_r(T) = \{\lambda \in \mathbb{C} : \lambda I - T \quad$ is injective, but not surjective and without dense range

So: elements of the point spectrum are the eigenvalues.

if $\lambda \in \sigma_p(T)$ then $\ker(\lambda I - T)$ is called the *Eigenspace*. Observe

Lemma 21.17. *$\ker(\lambda I - T)$ is invariant under T, in the following sense*

$$x \in \ker(\lambda I - T) \quad \Rightarrow Tx \in \ker(\lambda I - T).$$

Proof. If $x \in \ker(\lambda I - T)$ then

$$(\lambda I - T)Tx = T(\lambda I - T)x = 0$$

That is $Tx \in \ker(\lambda I - T)$. □

Example 21.18. Consider again the right-shift operator

$$S_+ : \ell^p \to \ell^p,$$

$$(x_1, x_2, \ldots,) \mapsto (0, x_1, x_2, \ldots).$$

We have

$$(\lambda I - S_+)x = (\lambda x_1, \lambda x_2 - x_1, \lambda x_3 - x_2, \ldots).$$

That is, for any $\lambda \in \mathbb{C}$, $(\lambda I - S_+)$ is injective, i.e. $\sigma_p(S_+) = \emptyset$.

Observe that $0 \in \sigma_r S_+$ since $\mathrm{im} S_+$ contains only sequeneces with first entry zero (i.e. is not dense in ℓ^p).

One can show that the operator S_+ is surjective if and only if $|\lambda| > 1$. Thus $\sigma(S_+) = B(0,1) \subset \mathbb{C}$.

The *spectral radius*

$$r(T) := \sup_{\lambda \in \sigma(T)} |\lambda|$$

We recall the von Neumann theorem:

Exercise 21.19. *Assume $T : X \to X$ is a linear bounded operator with $\|T\| < 1$*

Show that $I - T$ is invertible, i.e. there exists $(I-T)^{-1} : X \to X$ which is linear and bounded.

For this you can use the von Neumann sum

$$Sx := \lim_{N\to\infty} \sum_{k=0}^{N} T^k,$$

where $T^k = T \circ \ldots \circ T$. Show that S is a linear bounded operator, and it is indeed $(I-T)^{-1}$.

We have for the spectral radius

$$r(T) \leq \|T\|,$$

by the following

Lemma 21.20 (Spectral radius bound)**.** *If $\lambda \in \mathbb{C}$ and $|\lambda| > \|T\|$ then $\lambda \notin \sigma(T)$.*

Exercise 21.21. *Prove Lemma 21.20 – Hint: consider $(\frac{1}{\lambda}T - I)$ and use Exercise 21.19.*

Lemma 21.22. *Let X be a Banach space and $T : X \to X$ linear and bounded. Then $\sigma(T) \subset \mathbb{C}$ is a closed set in $\mathbb{C}$.*

Proof. We prove that $\rho(T) = \mathbb{C} \setminus \sigma(T)$ is open.

Let $\lambda_0 \in \rho(T)$. Since $(\lambda_0 I - T)$ is invertible we may write

$$\begin{aligned}
\lambda I - T =& (\lambda_0 I - T) - (\lambda_0 - \lambda) I \\
=& (\lambda_0 I - T)\left(I - (\lambda_0 - \lambda)(\lambda_0 I - T)^{-1}\right) \\
=& (\lambda_0 I - T)\left(I - \tilde{T}\right)
\end{aligned}$$

where we set

$$\tilde{T} := (\lambda_0 - \lambda)(\lambda_0 I - T)^{-1},$$

which is a bounded linear operator from $X \to X$.

By Exercise 21.19, if $\|\tilde{T}\| < 1$ then $I - \tilde{T}$ is invertible. On the other hand $\|\tilde{T}\| = \|\lambda_0 - \lambda\| \|(\lambda_0 I - T)^{-1}\| < 1$ whenever $|\lambda_0 - \lambda| \ll 1$. Thus

$$\lambda I - T = (\lambda_0 I - T)(I - \tilde{T})$$

is invertible for $\lambda \approx \lambda_0$, and thus $\lambda \in \rho(T)$. $\square$

Exercise 21.23. *Prove the following extension of Lemma 21.22:*

Let X be a Banach space and $T : X \to X$ linear and bounded. Take $\lambda_0 \in \rho(T) = \mathbb{C} \setminus \sigma(T)$. Then for any $\lambda \in \mathbb{C}$ with

$$|\lambda - \lambda_0| < \|T_{\lambda_0}\|^{-1}$$

we have that $\lambda I - T$ is bijective and

$$(\lambda I - T)^{-1} = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k T_{\lambda_0}^{k+1}.$$

Here $T_{\lambda_0} := (\lambda I - T)^{-1}$.

We can obtain the following structure

Theorem 21.24. *Let X be a Banach space and assume $T : X \to X$ is linear and bounded. Then*

(1) $\sigma(T)$ is compact
(2) $|\lambda| \leq \|T\|$ for all $\lambda \in \sigma(T)$
(3) If $X \neq \emptyset$ then $\sigma(T) \subset \mathbb{C}$ is nonempty.

Proof. The first two properties have been proven already: Lemma 21.20 implies that the spectrum $\sigma(T)$ is a bounded set, and Lemma 21.22 implies $\sigma(T)$ is closed.

It remains to prove (3).

Let $\xi \in L(X,X)^*$ a functional on the space $L(X,X)$ which consists of bounded linear mappings from X to X.

We consider $f_\xi : \rho(T) \subset \mathbb{C} \to \mathbb{C}$ given as

$$f_\xi(\lambda) := \xi[(\lambda I - T)^{-1}].$$

For $\lambda_0 \in \rho(T)$ whenever $|\lambda - \lambda_0| \ll 1$ by Exercise 21.23,

$$f_\xi(\lambda) = \xi[(\lambda I - T)^{-1}] = \sum_{k=0}^{\infty} (\lambda_0 - \lambda)^k \xi[(\lambda_0 I - T)^{-(k+1)}].$$

Here we have used the convergence of the series in $L(X,X)$ and the continuity of ξ.

Thus $f_\xi(\lambda)$ is a power series for $\mathbb{C} \ni \lambda \approx \lambda_0$. That is:

For any $\lambda_0 \in \rho(T)$ we know that $f_\xi(\lambda)$ is *holomorphic* for all $\lambda \approx \lambda_0$.

If $\sigma(T) = \emptyset$ then $\rho(T) = \mathbb{C}$, that is f is a power series everywhere, i.e. $f : \mathbb{C} \to \mathbb{C}$ is holomorphic.

We are now going to use *Liouville theorem* (bounded holomorphic functions in $\mathbb{C}$ are constant):

Observe that if $\lambda > 2\|T\|$ then, by von Neumann series Exercise 21.19,

$$(\lambda I - T)^{-1} = \lambda^{-1} \sum_{k=0}^{\infty} (\lambda^{-1} T)^k,$$

since

$$\|\lambda^{-1} T\|_{L(X,X)} \leq \frac{1}{2}.$$

That is,

$$|f_\xi(\lambda)| \leq \|\xi\|_{L(X,X)^*} \frac{1}{\lambda} \|\sum_{k=0}^{\infty} (\lambda^{-1} T)^k\|_{L(X,X)} \leq \frac{2}{\lambda} \|\xi\|_{L(X,X)^*}.$$

That is, f_ξ decays at infinity, since it is holomorphic it is also continuous: that is f is a globally bounded holomorphic function. By Liouville theorem, f_ξ must be constant. Since $f_\xi = 0$ at infinity we see that $f_\xi \equiv 0$.

That is for all $\xi \in L(X,X)^*$ and all $\lambda \in \mathbb{C}$ we have

$$\xi[(\lambda I - T)^{-1}] = 0.$$

But then (by Hahn-Banach, e.g. Corollary 10.14) $(\lambda I - T)^{-1} = 0$ which is impossible because $(\lambda I - T)^{-1}$ must be invertible (as the inverse of a function). Contradiction, and we can conclude. □

We recall the notion of spectral radius

$$r(T) := \sup_{\lambda \in \sigma(T)} |\lambda|$$

Now we get all algebraic:

Let p be any (complex) polynomial of degree n,

$$p(z) = \sum_{k=0}^{n} a_k z^k.$$

Then we can define

$$p(T) := \sum_{k=0}^{n} a_k T^k.$$

Here $T^k = T \circ \ldots T$ k times and $T^0 = I$.

Then $p(T) : X \to X$ is still a bounded linear operator as a composition of bounded and linear operators.

We have the following curious relation

Theorem 21.25 (Spectral polynomial theorem). *Let X be a Banach space and $T \in L(X,X)$ linear and bounded, and let p be a polynomial as above.*

Then

$$\sigma(p(T)) = p(\sigma(T)) \equiv \{p(\lambda) : \quad \lambda \in \sigma(T)\}.$$

Proof. If p is constant, then $p(T) = a_0 T$, and the above statement is obvious.

Assume now p has degree at least 1.

Assume $\lambda \in \sigma(p(T))$. We show that $\lambda \in p(\sigma(T))$, i.e. there exists $\mu \in \sigma(T)$ such that $\lambda = p(\mu)$.

Set

$$q(z) = p(z) - \lambda.$$

This is a polynomial of degree n. The fundamental theorem of algebra gives us n roots $\mu_1, \ldots, \mu_n \in \mathbb{C}$, and we write for some $c \in \mathbb{C} \setminus \{0\}$

$$q(z) = c\Pi_{k=1}^n (z - \mu_k).$$

Since $\lambda \in \sigma(p(T))$ we know that $p(T) - \lambda I$ is not bijective. The above formula implies that there must be some $\bar{k} \in \{1, \ldots, n\}$ such that $T - \mu_{\bar{k}} I$ cannot be bijective, so this $\mu_{\bar{k}} \in \sigma(T)$.

We then have

$$\lambda = \lambda + \underbrace{q(\mu_{\bar{k}})}_{=0} = p(\mu_{\bar{k}}).$$

That is, $\lambda \in p(\sigma(T))$.

Now assume $\lambda \in \sigma(T)$. We show that $p(\lambda) \in \sigma(p(T))$.

Set

$$q(z) := p(z) - p(\lambda).$$

Then λ is a root of q, and thus

$$q(z) = (z - \lambda) r(z),$$

for some polynomial r of degree at most $n - 1$.

Then

$$p(T) - p(\lambda) I = q(T).$$

Since $\lambda \in \sigma(T)$, we have $T - \lambda I$ is not bijective.

We need to conclude that $q(T)$ is not bijective. Assume to the contrary it is, then by the following exercise and since

$$q(T) = (T - \lambda I) r(T) = r(T)(T - \lambda I),$$

we have that $r(T)$ must be bijective (since $T - \lambda I$ is not bijective). But then

$$q(T) r(T)^{-1} = (T - \lambda I),$$

that is $T - \lambda I$ is bijective, a contradiction. Thus $q(T) = p(T) - p(\lambda)$ is not bijective, and thus $p(\lambda) \in \sigma(p(T))$. $\square$

Exercise. *Let X be a Banach space and let $S, T : X \to Y$ be bounded, linear operators. Assume that neither S, T are bijective.*

(1) Give an example of X, S, T as above such that ST is bijective
(2) Show that $S \circ T$ or $T \circ S$ are not bijective.

Exercise 21.26. *Let X be a Banach space and $T : X \to X$ linear and bounded. Show that*

$$\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} \|T^n\|_{L(X)}^{\frac{1}{n}}.$$

Hint: $\|T^2\| \leq \|T\|^2$.

Theorem 21.27 (Spectral radius theorem)**.** *Let $X \neq \emptyset$ be a Banach space and $T : X \to X$ linear and bounded. Then the spectal radius $r(T)$ can be computed by*

$$r(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\in\mathbb{N}} \|T^n\|_{L(X)}^{\frac{1}{n}}.$$

Proof. We prove first the lower bound

$$r(T) \leq \inf_{n\in\mathbb{N}} \|T^n\|_{L(X)}^{\frac{1}{n}}.$$

Assume $\lambda \in \sigma(T)$ (observe that by Theorem 21.24, $\sigma(T) \neq \emptyset$).

By the *spectral polynomial theorem*, Theorem 21.25, we have $\lambda^n \in \sigma(T^n)$. By Lemma 21.20 this implies $|\lambda^n| \leq \|T^n\|$. That is

$$|\lambda| \leq \|T^n\|^{\frac{1}{n}}.$$

Take the supremum over all $\lambda \in \sigma(T)$ then we have shown

$$r(T) \leq \|T^n\|^{\frac{1}{n}} \quad \forall n \in \mathbb{N}.$$

That is

$$r(T) \leq \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}}.$$

This establishes the lower bound.

Now we prove the upper bound. Recall the function

$$f_\xi : \rho(T) = \mathbb{C} \setminus \sigma(T) \to \mathbb{C}$$

given by

$$f_\xi(\lambda) := \xi[(\lambda I - T)^{-1}]$$

where $\xi \in L(X,X)^*$ was any fixed functional, from the proof of Theorem 21.24. We had shown already that f is holomorphic around any point $\lambda \in \rho(T) = \mathbb{C} \setminus \sigma(T)$.

In particular, f_ξ is holomorphic in $\{|\lambda| > r(T)\}$, and we have for $|\lambda| > \|T\|$ (by von Neumann series, Exercise 21.19) the *Laurent series*

$$f_\xi(\lambda) = \sum_{k=0}^{\infty} \lambda^{-k-1}\xi[T^k].$$

This Laurent series must be valid in the whole annulus $\{|\lambda| > r(T)\}$. That is the above series converges, in particular

$$\left|\lambda^{-k-1}\xi[T^k]\right| \xrightarrow{k\to\infty} 0.$$

This holds for any ξ in $L(X,X)^*$ so we have *weak convergence* (see Section 12) for any $|\lambda| > r(T)$,

$$\lambda^{-k-1}T^k \hookrightarrow 0 \quad \text{weakly in } L(X,X).$$

Weakly convergent sequences are bounded, Theorem 12.17, so we have

$$\sup_{k\in\mathbb{N}} \|\lambda^{-k-1}T^k\|_{L(X,X)} < \infty.$$

That is for any $|\lambda| > r(T)$ ther must be some $C = C(\lambda) > 0$ such that

$$\|T^k\|^{\frac{1}{k}} \leq \left(C|\lambda|^{k+1}\right)^{\frac{1}{k}} \xrightarrow{k\to\infty} |\lambda|$$

That is for any $|\lambda| > r(T)$

$$\limsup_{k\to\infty} \|T^k\|^{\frac{1}{k}} \leq |\lambda|,$$

we conclude that

$$\limsup_{k\to\infty} \|T^k\|^{\frac{1}{k}} \leq r(T).$$

Combining this with the lower bound and Exercise 21.26 we conclude. □

Next we try to recover properties of eigenvalues, the first one is that eigenspaces are linearly independent:

Theorem 21.28. *Let $T : X \to X$ be linear and bounded. Then eigenvectors corresponding to distinct eigenvalues are linearly independent.*

Proof. We argue by induction on the number of eigenvectors N. If we have only one eigenvector there is nothing to show.

So now we assume that we have already shown that if we have N eigenvectors to distinct eigenvalues, then they are linearly independent.

Let now

$$(x_i)_{i=1}^{N+1} \quad \in X \setminus \{0\}$$

be eigenvectors of T to eigenvalues $\lambda_i \in \mathbb{C}$, i.e. $Tx_i = \lambda_i x_i$, where $(\lambda_i)_{i=1}^{N+1}$ are pairwise distrinct, $\lambda_i \neq \lambda_j$ for all $i \neq j$.

Assume that

$$0 = \sum_{i=1}^{N+1} \mu_i x_i = 0.$$

We can apply T to that sum and obtain that

$$0 = \sum_{i=1}^{N+1} \mu_i \lambda_i x_i = 0.$$

One of the eigenvalues λ_i is nonzero, w.l.o.g. $\lambda_{N+1} \neq 0$. Then we have two equations

$$\begin{cases} 0 = \sum_{i=1}^{N+1} \mu_i x_i = 0 \\ 0 = \sum_{i=1}^{N+1} \frac{\lambda_i}{\lambda_{N+1}} \mu_i x_i = 0 \end{cases}$$

Subtracting the last from the first equation we find that

$$0 = \sum_{i=1}^{N} \left(1 - \frac{\lambda_i}{\lambda_{N+1}}\right) \mu_i x_i = 0$$

Since the first N eigenvalues must be linearly independent by induction hypothesis, we conclude that

$$\left(1 - \frac{\lambda_i}{\lambda_{N+1}}\right) \mu_i = 0 \quad i = 1, \ldots, N$$

since $\lambda_i \neq \lambda_{N+1}$ for $i = 1, \ldots, N$ we conclude

$$\mu_i = 0 \quad i = 1, \ldots, N.$$

We conclude that also $\mu_{N+1} = 0$ (since $x_{N+1} \neq 0$) so we have that $\mu_i = 0$ for $i = 1, \ldots, N+1$, and thus $(x_i)_{i=1}^{N+1}$ must have been linearly independent. □

21.6. **Spectrum for compact operators.** When looking at *compact* operators, we can say more. First of all: any *nonzero* λ in the spectrum is an eigenvalue. And if the dimension is infinite, 0 is in the spectrum (but it may or may not be an eigenvalue)

Theorem 21.29. *Assume that $T : X \to X$ is a linear bounded and compact operator and X is a Banach space. Then every nonzero element in $\lambda \in \sigma(T) \setminus \{0\}$ is an eigenvalue. Also if $\dim(X) = \infty$ then $0 \in \sigma(T)$.*

Proof. We can apply the Fredholm alternative, Theorem 21.13. If $\lambda \neq 0$ is not an eigenvalue (i.e. there is no nontrivial solution to the homogeneous equation $(\lambda I - K)v = 0$), then by Fedholm alternative $(\lambda I - K)$ must be surjective (and since λ is not an eigenvalue it is injective). Thus $\lambda I - K$ is bijective, thus invertible – meaning that $\lambda \notin \sigma(T)$.

Also, since T is compact, in view of Corollary 20.18 (which is a consequence of open mapping theorem and Riesz Lemma), T is not surjective. Meaning $0 \in \sigma(T)$ since $T - 0I$ is not invertible. □

Lemma 21.30. *Let $T : X \to X$ be compact and $\lambda \neq 0$ an eigenvalue. Then the eigenspace*

$$\ker(\lambda I - T)$$

is finite dimensional.

Proof. Since $\lambda \neq 0$ we can consider $\lambda I - T = \lambda(I - \frac{1}{\lambda} T)$. $\frac{1}{\lambda} T$ is still compact, so w.l.o.g. $\lambda = 1$.

Set $Y := \ker(I - T)$. Then

$$Ty = y \quad \forall y \in Y.$$

Thus $T\Big|_Y : Y \to Y$ is the identity, in particular $T\Big|_Y : Y \to Y$ is surjective. Since on the other hand $T\Big|_Y$ is still compact, by Corollary 20.18, Y must be finite dimensional. $\square$

Theorem 21.31. *Assume that $T : X \to X$ is a linear bounded and compact operator and X is a Banach space. Then eigenvalues form a finite or countable set, i.e.*

$$\sigma(T) = \{0\} \cup \{\lambda_i\}_{i=1}^N \quad \text{if } X \text{ is infinite dimensional},$$

or

$$\sigma(T) = \{\lambda_i\}_{i=1}^N \quad \text{if } X \text{ is finite dimensional},$$

where $N \in \{0, 1, \dots, \} \cup \{\infty\}$ and each $\lambda_i \in \mathbb{C} \setminus \{0\}$ is an eigenvalue. If $N = \infty$ then we can order

$$|\lambda_1| \geq |\lambda_2| \geq \dots$$

and then have $\lim_{i\to\infty} \lambda_i = 0$.

Proof. It suffices to prove that for any $r > 0$ the number of eigenvalues satisfying $|\lambda| \geq r$ is finite. If this was not true, then there'd be distinct eigenvalues

$$\lambda_i, \quad |\lambda_i| \geq r, i = 1, 2, 3, \dots$$

and corresponding eigenvectors

$$Tx_i = \lambda_i x_i, \quad \|x_i\| = 1.$$

For $n \in \mathbb{N}$ we consider the (closed) linear spaces

$$H_n = \mathrm{span}\{x_1, \dots, x_n\}.$$

Since each x_i belongs is eigenvector to a different eigenvalue λ_i, we have that $\dim H_n = n$, Theorem 21.28.

By the Riesz lemma, Lemma 7.18, we can find $w_n \in H_n$ such that

$$\|w_n\| = 1, \|w_n - y\| \geq \frac{1}{2} \quad \text{for any } y \in H_{n-1}.$$

We can represent $w_n = a_n x_n + y_{n-1}$ for some $y_{n-1} \in H_{n-1}$, $a_n \in \mathbb{R}$. Thus for $k < n$, $Tw_k \in H_k \subset H_{n-1}$ and hence

$$\|Tw_n - Tw_k\| = \|a_n \lambda_n x_n + Ty_{n-1} - Tw_k\|$$
$$= |\lambda_n| \, \| \underbrace{a_n x_n + y_{n-1}}_{w_n} - \underbrace{(y_{n-1} - \lambda_n^{-1}(Ty_{n-1} - Tw_k))}_{\in H_{n-1}} \| \geq \frac{r}{2}.$$

Therefore the sequence $\{Tw_n\}$ has no convergent subsequence which contradicts compactness of T. $\square$

21.7. **Spectral theory in Hilbert spaces.** We now restrict our attention further to Hilbert spaces.

We work in complex Hilbert spaces, so in particular

$$\langle v, w\rangle \in \mathbb{C}$$

and

$$\langle v, w\rangle = \overline{\langle w, v\rangle},$$

and (this is a choice)

$$\langle \lambda u + \mu v, w\rangle = \lambda\langle u, w\rangle + \mu\langle v, w\rangle,$$

and thus

$$\langle w, \lambda u + \mu v\rangle = \overline{\lambda}\langle w, u\rangle + \overline{\mu}\langle w, v\rangle.$$

If $T \in \mathcal{L}(H, H)$ where H is a Hilbert space, then $T^* : H \to H$. In particular we can talk about what it means to say $T = T^*$. Observe for (real) matrices T we have $T^* = T$ if $T^t = T$. For complex matrices these are unitary matrices $\bar{U}^t = U$.

Definition 21.32. Let H be a Hilbert space. A linear operator $T : H \to H$ is said to be *self-adjoint* if

$$\langle Tx, y\rangle = \langle x, Ty\rangle \quad \forall x, y \in H.$$

Observe that in view of Proposition 20.27 any such T is a bounded operator.

Theorem 21.33. *Let $T \in \mathcal{L}(H, H)$ be self-adjoint then*

1. *$\langle Tx, x\rangle$ is real*
2. *Eigenvalues are real*
3. *Eigenspace of T corresponding to distrinct eigenvalues are orthogonal to each other*[46]

Proof. Observe we use here the complex scalar product, i.e. $\langle Tx, x\rangle \in \mathbb{C}$ in general. However we have

$$\langle Tx, x\rangle = \langle x, Tx\rangle = \overline{\langle Tx, x\rangle},$$

and thus $\langle Tx, x\rangle$ must be real.

In particular if λ is an eigenvalue of T and x an eigenvector, then

$$\mathbb{R} \ni \langle Tx, x\rangle = \lambda\langle x, x\rangle = \lambda\|x\|^2$$

Thus $\lambda \in \mathbb{R}$ as well.

Assume that $\lambda_1 \neq \lambda_2$ are two eigenvalues with corresponding eigenvectors v_1 and v_2. Then

$$\lambda_1\langle v_1, v_2\rangle = \langle Tv_1, v_2\rangle = \langle v_1, Tv_2\rangle = \bar{\lambda}_2\langle v_1, v_2\rangle.$$

Since $\lambda_1, \lambda_2 \in \mathbb{R}$ this implies

$$(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0,$$

i.e. $v_1 \perp v_2$ (since $\lambda_1 \neq \lambda_2$). $\square$

[46] from Theorem 21.28 we already know linear independence

We already know that any $\sigma(T)$ is bounded by $\|T\|$, Lemma 21.20. Here we can obtain the maximum value

Theorem 21.34. *Let $T \in \mathcal{L}(H,H)$ be self-adjoint and compact, the $\|T\|$ or $-\|T\|$ is an eigenvalue of T.*

Proof. Let $x_n \in H$ be a sequence such that $\|x_n\| = 1$, $\|Tx_n\| \to \|T\|$. Since T is compact, we may assume that $Tx_n \xrightarrow{n\to\infty} y \in H$, and thus $\|y\| = \|T\|$.

Moreover we have $T^2 x_n \xrightarrow{n\to\infty} Ty$. Then

$$\|Ty\| = \lim_{n\to\infty} \|T^2 x_n\| \overset{C.S.}{\geq} \lim_{n\to\infty} \langle T^2 x_n, x_n\rangle \overset{T^*=T}{=} \lim_{n\to\infty} \langle Tx_n, Tx_n\rangle = \|T\|^2.$$

Repeating this argument

$$\|T^2 y\|\|y\| \geq \langle T^2 y, y\rangle = \|Ty\|^2 \geq \|T\|^4 = \|T\|^2\|y\|^2 \geq \|T^2 y\|\|y\|.$$

Thus we actually have an equality

$$\langle T^2 y, y\rangle = \|T^2 y\|\|y\|.$$

This is equality in the Cauchy-Schwartz inequality – which means that $T^2 y$ and y must be collinear, i.e.

$$T^2 y = \mu y$$

for some $\mu \in \mathbb{C}$ (but since T^2 is still self-adjoint, we already know that $\mu \in \mathbb{R}$). Also observe that we already know $T^2 y \neq 0$ from the above estimates.

We have

$$\mu = \frac{\langle T^2 y, y\rangle}{\langle y, y\rangle} = \frac{\|T\|^4}{\|T\|^2} = \|T\|^2.$$

Let $x = y + \|T\|^{-1} Ty$. If $x = 0$ then $Ty = -\|T\| y$ and hence $-\|T\|$ is an eigenvalue of T. If $x \neq 0$ then

$$Tx = Ty + \|T\|^{-1} T^2 y = Ty + \|T\|^{-1}\|T\|^2 y = Ty + \|T\| y = \|T\| x.$$

That is $\|T\|$ is an eigenvalue. □

The following is the infinite dimensional version of the spectral theorem of Algebra (saying that we can diagonalize symmetric matrices)

Theorem 21.35 (Spectral Theorem)**.** *Let $T : H \to H$ be compact and selfadjoint linear map, where H is a Hilbert space.*

Then there exists a finite or countable sequence of real eigenvalues $(\lambda_i)_{i=1}^N \in \mathbb{R}$, $|\lambda_1| > |\lambda_2| \geq \ldots$, for some $N \in \{0,1,2,\ldots\} \cup \{\infty\}$. If $N = \infty$ then $\lim_{i\to\infty} \lambda_i = 0$.

(1) For each eigenvalue λ_i, its corresponding eigenspace is finite dimensional

$$\dim \ker(\lambda_i I - T) < \infty.$$

(2) *For two distinct eigenvalues the eigenspaces are perpendicular*
$$\ker(\lambda_i I - T) \perp \ker(\lambda_j I - T) \quad \textit{whenever } i \neq j.$$

(3) *And, we can decompose H into $\ker T$ and the eigenspaces, namely*
$$H = \ker T \oplus \bigoplus_{i=1}^{N} \ker(\lambda_i I - T),$$
in the following sense. For any vector $x \in H$ there exists exactly one $y \in \ker T$, $y_i \in \ker(\lambda_i I - T)$ such that
$$x = y + \sum_{i=1}^{N} y_i \quad \textit{in } H$$
Moreover y and y_i are all mutually perpendicular.

In particular if $\ker T = 0$ we can form a orthonormal basis of eigenvectors of X, i.e. there exists a countable sequence $(o_k)_{k=0}^{M}$ where $M \geq N$, $\|o_k\| = 1$, $o_k \perp o_j$ for $k \neq j$, each o_k is an eigenvector to some λ_j, and for each $x \in H$,
$$x = \sum_{k=1}^{M} \langle x, o_k \rangle \, o_k \quad \textit{convergence in } H$$

Proof. We obtain the sequence of eigenvalues $\lambda_i \neq 0$ from Theorem 21.31.

(1) This is true since T is compact by Lemma 21.30.
(2) This follows from Theorem 21.33
(3) Set
$$Y := \bigoplus_{i=1}^{N} \ker(\lambda_i I - T),$$
i.e. for each $y \in H$,
$$y \in Y \quad :\Leftrightarrow y = \sum_{i=1}^{N} y_i,$$
where $y_i \in \ker(\lambda_i - T)$. Observe that the right-hand side assumes convergence in H! Clearly, Y is a linear space. We can also check it is a closed linear space: let $z_k \in Y$, $\bar{z} = \lim_{k\to\infty} z_k \in H$. Then
$$z_k = \sum_{i=1}^{N} y_{i;k},$$
and thus
$$z_k - z_\ell = \sum_{i=1}^{N} (y_{i;k} - y_{i;\ell}).$$
Since for $i \neq j$,
$$(y_{i;k} - y_{i;\ell}) \in \ker(\lambda_i I - T) \perp \ker(\lambda_j I - T) \ni (y_{j;k} - y_{j;\ell})$$

we have

$$\|z_k - z_\ell\|_H^2 = \sum_{i=1}^N \|y_{i;k} - y_{i;\ell}\|_H^2. \tag{21.1}$$

So in particular

$$\|z_k - z_\ell\|_H < \varepsilon \Rightarrow \sup_i \|y_{i;k} - y_{i;\ell}\|_H < \varepsilon.$$

This implies that for each fixed i, the sequence $(y_{i;k})_{k\in\mathbb{N}}$ is a Cauchy sequence, thus there exists a limit $\bar{z}_i := \lim_{k\to\infty} y_{i;k}$. Since $y_{i;k} \in \ker(\lambda_i I - T)$ (which is a closed space) we have $z_i \in \ker(\lambda_i I - T)$.

For each $L \in N$ and any $k \in \mathbb{N}$ we then have

$$\sum_{i=1}^L \|y_{i;k} - \bar{y}_i\|_H^2 \lesssim \sum_{i=1}^L \|y_{i;k} - y_{i;\ell}\|_H^2 + \sum_{i=1}^L \|y_{i;\ell} - \bar{y}_i\|_H^2$$

Thus,

$$\sum_{i=1}^L \|y_{i;k} - \bar{y}_i\|_H^2 \lesssim \|z_k - z_\ell\|_H + \sum_{i=1}^L \|y_{i;\ell} - \bar{y}_i\|_H^2.$$

So for each $\varepsilon > 0$ there exist some $K \in \mathbb{N}$ such that for all $k, \ell > K$,

$$\sum_{i=1}^L \|y_{i;k} - \bar{y}_i\|_H^2 \lesssim \varepsilon + \sum_{i=1}^L \|y_{i;\ell} - \bar{y}_i\|_H^2.$$

Taking the limit in $\ell \to \infty$ we have

$$\sum_{i=1}^L \|y_{i;k} - \bar{y}_i\|_H^2 \lesssim \varepsilon.$$

This holds independently of L, and thus for any $\varepsilon > 0$ there exist $K \in \mathbb{N}$ such that for $k > K$

$$\sum_{i=1}^\infty \|y_{i;k} - \bar{y}_i\|_H^2 \lesssim \varepsilon.$$

In particular we have $\sum_{i=1}^\infty \bar{y}_i \in H$ and for all $k > K$,

$$\left\| z_k - \sum_{i=1}^\infty \bar{y}_i \right\|_H^2 \lesssim \varepsilon,$$

that is,

$$\lim_{k\to\infty} \left\| z_k - \sum_{i=1}^\infty \bar{y}_i \right\|_H^2 = 0.$$

Consequently,

$$\left\| z - \sum_{i=1}^\infty \bar{y}_i \right\|^2 \lesssim \|z - z_k\|^2 + \left\| z_k - \sum_{i=1}^\infty \bar{y}_i \right\|^2 \xrightarrow{k\to\infty} 0.$$

Thus $z = \sum_{i=1}^\infty \bar{y}_i \in Y$ – i.e. Y is closed.

Next we are going to show that

$$\ker T = Y^\perp.$$

For $\subseteq$ we see that if $\bar{y} \in H$ such that $Ty = 0$, then

$$\begin{aligned}\langle \sum_{i=1}^N y_i, y\rangle =& \sum_{i=1}^N \langle y_i, y\rangle \overset{Ty_i=\lambda y_i}{=} \sum_{i=1}^N \underbrace{\frac{1}{\lambda_i}}_{\neq\infty} \langle {\color{red}T} y_i, y\rangle\\
=& \sum_{i=1}^N \frac{1}{\lambda_i}\langle y_i, Ty\rangle\\
=& \sum_{i=1}^N \frac{1}{\lambda_i}\langle y_i, 0\rangle\\
=& 0\end{aligned}$$

So we have for any $y \in \ker T$ that $y \in Y^\perp$.

For $\supseteq$ let $z \in Y^\perp$. We claim that $Tz \in Y^\perp$. For this take any $y \in Y$. Then, by definition of Y, for some $y_i \in \ker(\lambda_i I - T)$ we have

$$y = \sum_{i=1}^N y_i,$$

and thus

$$Ty = T\left(\lim_{k\to\infty} \sum_{i=1}^{\min\{N,i\}} y_i\right) = \lim_{k\to\infty} \sum_{i=1}^{\min\{N,i\}} Ty_i = \lim_{k\to\infty} \sum_{i=1}^{\min\{N,i\}} \underbrace{\lambda_i y_i}_{\in\ker(\lambda_i I - T)} \in Y.$$

Consequently,

$$\langle Tz, y\rangle = \langle \underbrace{z}_{\in Y^\perp}, \underbrace{Ty}_{\in Y}\rangle = 0.$$

That is $T\Big|_{Y^\perp} : Y^\perp \to Y^\perp$ is a linear compact, self-adjoint operator. Observe that the only eigenvalue $T\Big|_{Y^\perp}$ could have is 0 (because any other eigenvalue must lie in Y). However in view of Theorem 21.34, $\left\|T\Big|_{Y^\perp}\right\|$ or $-\left\|T\Big|_{Y^\perp}\right\|$ must be an eigenvalue of $T\Big|_{Y^\perp}$ – thus $0 = \left\|T\Big|_{Y^\perp}\right\|$ and we conclude that $Y^\perp \subset \ker T$.

We have thus established

$$\ker T = Y^\perp.$$

By Lemma 19.13, using that Y is closed we have

$$H = \ker T \oplus Y.$$

We can conclude.

□

Exercise 21.36. *Let* $(\lambda_i)_{i=1}^\infty \subset \mathbb{C}$ *be any sequence with* $\lim_{i\to\infty} \lambda_i = 0$.

Set

$$T : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$$

be defined as follows

$$T(x_1, x_2, \ldots) = (\lambda_1 x_1, \lambda x_2, \ldots,).$$

Show that

- T *is compact.*
- T *is self-adjoint on* $\ell^2(\mathbb{N})$ *if and only if* $\lambda_i \in \mathbb{R}$ *for all* $i \in \mathbb{N}$
- *Show that the eigenvalues of* T *are exactly* $\lambda_1, \ldots, .$
- *If* $\lambda_i \neq 0$ *for all* i *show that* T *is injective, i.e.* $\ker T \neq 0$.
- *Given an example to show that* T *is not surjective (which according to Corollary 20.18 it can't be).*

21.8. **Spectral theorem for the Laplace-Operator.** We want to study the eigenvalues of the Laplace operator Δ. This is not so obvious, because for what space would we have

$$\Delta : X \to X?$$

We could take $X = L^2$, and consider $\Delta : D(\Delta) \to L^2$ as densily defined operator by taking $D(\Delta) = W^{2,2}$. But this is very messy (also then Δ is not compact...)

The principle idea is that (at least formally) if we want to consider eigenvalues, e.g.

$$Tu = \lambda u$$

then if we can apply T^{-1} (i.e. if it makes sense) this is as good as

$$u = \lambda T^{-1} u.$$

That is (formally) $\frac{1}{\lambda}$ is an eigenvalue of the inverse operator if λ is an eigenvalue of T and vice versa.

To make this precise, we need to precisely write what we want. We say that $\lambda \in \mathbb{C}$ is an eigenvalue for the (Dirichlet-)Laplace problem in Ω (for simplicity: open, bounded set with smooth boundary) if there exists a solution $u \neq 0$ of

$$\begin{cases} -\Delta u = \lambda u & \text{in } \Omega \\ u = 0 & \text{in } \Omega \end{cases}$$

Observe that if we change Ω, or change the boundary data then the Eigenvalue of Δ will likely change[47].

[47]this is related to the famous 1966 article "Can One Hear the Shape of a Drum?" article by Mark Kac in the American Mathematical Monthly

So if we want to consider the inverse operator for Δ (depending on Ω of course) then we want to consider the solution operator T that maps $f \in L^2(\Omega)$ to $u \in W^{1,2}_0(\Omega)$ solving

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{in } \Omega \end{cases}$$

From existence theory (we have considered various arguments) $T : L^2(\Omega) \to W^{1,2}_0(\Omega) \subset L^2(\Omega)$ is well-defined linear operator and we have

$$\|Tf\|_{W^{1,2}(\Omega)} \leq \|f\|_{L^2(\Omega)}.$$

(Actually from regularity theory we know more: we know the same estimate for $W^{2,2}(\Omega)$ on the left-hand side, but that does not matter for us, so we are content with the $W^{1,2}$-estimate).

Now if

$$\begin{cases} -\Delta u = \lambda u & \text{in } \Omega \\ u = 0 & \text{in } \Omega \end{cases}$$

Then this implies $T\lambda u = u$ and thus $\frac{1}{\lambda} Tu = u$ – and vice versa. So Eigenvalues of T are as good as discussing eigenvalues of Δ.

Moreover $T : L^2(\Omega) \to L^2(\Omega)$ is compact. Indeed, let $(f_k)_{k\in\mathbb{N}}$ be a bounded sequence in $L^2(\Omega)$,

$$\sup_k \|f_k\|_{L^2(\Omega)} < \infty.$$

Then by the above estimate

$$\sup_k \|Tf_k\|_{W^{1,2}(\Omega)} < \infty.$$

By Rellich's theorem, Theorem 13.35 there exists a subsequence $(f_{k_i})_{i\in\mathbb{N}}$ such that f_{k_i} converges strongly to some limit f with respect to the $L^2(\Omega)$-norm (here we use the boundary regularity of $\partial\Omega$) – thus $T : L^2(\Omega) \to L^2(\Omega)$ is indeed compact.

Proposition 21.37. *The operator $T : L^2(\Omega) \to L^2(\Omega)$ is a self-adjoint, compact operator.*

Proof. Compactness we have discussed. As for self-adjointness, we need to show

$$\langle Tf, g\rangle = \langle f, Tg\rangle \quad \forall f, g \in L^2(\Omega). \tag{21.2}$$

Set $u := Tf$ and $v := Tg$. Then $u \in W^{1,2}_0(\Omega)$ and $-\Delta u = f$, that is (the distributional definition)

$$\int \nabla u \nabla \varphi = \int f \ \varphi \quad \forall \varphi \in C^\infty_c(\Omega).$$

By density we can use the above equation also for $\varphi \in W^{1,2}_0(\Omega)$. Indeed, let $\varphi_k \in C^\infty_c(\Omega)$, $\varphi_k \xrightarrow{k\to\infty} \varphi$ in $W^{1,2}_0(\Omega)$ (this smooth sequence exists by definition of $W^{1,2}_0$!). Then we have

$$\int \nabla u \nabla \varphi \xleftarrow{k\to\infty} \int \nabla u \nabla \varphi_k = \int f \ \varphi \quad \forall \varphi_k \xrightarrow{k\to\infty} \int f \varphi$$

Thus we actually have

$$\int \nabla u \nabla \varphi = \int f \ \varphi \quad \forall \varphi \in W_0^{1,2}(\Omega).$$

In particular we can take $\varphi := v$ and then have

$$\int \nabla u \nabla v = \int f \ v = \langle f, Tg \rangle.$$

On the other hand, using the equation $\Delta v = g$, tested with $u \in W_0^{1,2}(\Omega)$ we also have

$$\int \nabla u \nabla v = \int ug = \langle Tf, g \rangle.$$

So we have established (21.2) and T is self-adjoint. □

We can now apply the spectral theorem to T (and thus $-\Delta$) By Theorem 21.33 all eigenvalues of T (and thus of Δ) must be real numbers, and they form a (finite or countable) sequence $(\mu_i)_{i=1}^N$, where $N \in \mathbb{N} \cup \{+\infty\}$.

Lemma 21.38 ($-\Delta$ is a positive operator)**.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded.*

Then all eigenvalues λ of $-\Delta$ are strictly positive.

Proof. Assume that $u \in W_0^{1,2}(\Omega)$ solves

$$-\Delta u = \lambda u \quad \text{in } \Omega,$$

which means

$$\int \nabla u \nabla \varphi = \lambda \int u\varphi \quad \forall \varphi \in W_0^{1,2}(\Omega).$$

Testing this equation with $\varphi := u \in W_0^{1,2}(\Omega)$ itself we have

$$\int_\Omega |\nabla u|^2 = \lambda \int_\Omega |u|^2.$$

If $\lambda \leq 0$ this implies that $|\nabla u| \equiv 0$, that is $u \equiv \mathit{const}$ in each connected component of Ω, and thus since $u \in W_0^{1,2}(\Omega)$, $u \equiv 0$. But then u is not an eigenfunction to λ.

That is, whenever u is an eigenfunction of λ we have $\lambda > 0$. □

So when we apply the spectral theorem Theorem 21.35, we find that

$$L^2(\Omega) = \ker T \oplus \bigoplus_{i=1}^N \ker(T - \mu_i I)$$

Since $\ker T = 0$ (0 is not an eigenvalue by Lemma 21.38) and the dimension of the Eigenspaces $\ker(T - \mu_i I)$ is finite, Lemma 21.30, we must have $N = \infty$ and thus there are must be a decreasing sequence $\mu_1 > \mu_2 > \ldots > 0$ such that $\lim_{i\to\infty} \mu_i = 0$.

Since $T = (-\Delta)^{-1}$ we have for $\lambda_i := \frac{1}{\mu_i}$ the following results

Theorem 21.39. *Let $\Omega \subset \mathbb{R}^n$ be open, bounded, and with smooth boundary $\partial\Omega \in C^\infty$.*

(1) *The eigenvalues of the Laplace operator* $-\Delta : W^{1,2}_0(\Omega) \to L^2(\Omega)$ *are positive and form an increasing sequence*

$$\lambda_1 < \lambda_2 < \lambda_3 < \dots, \quad \lim_{i\to\infty} \lambda_i = \infty.$$

(2) *The dimensions of the corresponding eigenspaces are finite,* $\dim E_{\lambda_i} < \infty$*,* $i = 1, 2, 3 \dots$

(3) *We have*

$$L^2(\Omega) = \bigoplus_{i=1}^{\infty} E_{\lambda_i}$$

and hence we may choose an orthonormal basis of $L^2(\Omega)$ *consisting of eigenfunctions* $(u_k)_{k\in\mathbb{N}}$ *of* $-\Delta$*.*

Observe that by regularity theory all eigenvalues must be $C^\infty(\overline{\Omega})$*.*

In particular we may expand any $f \in L^2(\Omega)$ into

$$f = \sum_{k=1}^{\infty} u_k \langle u_k, f\rangle.$$

Then the action of the Laplacian $-\Delta$ on f is (at least formally)

$$-\Delta f = \sum_{k=1}^{\infty} -\Delta u_k \langle u_k, f\rangle = \sum_{k=1}^{\infty} \lambda_k u_k \langle u_k, f\rangle.$$

Observe that in general the series on the right-hand side has no reason to converge in $L^2(\Omega)$ – since $-\Delta f$ may not belong to L^2!

Observe also that in general u_k and λ_k depend on the shape of Ω!.

We can use this spectral decomposition to write a power of $(-\Delta)$. Using that $\lambda_k > 0$, we can write

$$(-\Delta)^s f := \sum_{k=1}^{\infty} (\lambda_k)^s u_k \langle u_k, f\rangle.$$

This is called the *spectral fractional Laplacian*.

Example 21.40. Let $n = 1$ and $\Omega = (0,1)$. Then $-\Delta = -\frac{d^2}{dx^2}$ and one can check

$$w_n(x) := \sqrt{2}\sin(n\pi x), n \in \mathbb{N}$$

are eigenfunctions of $-\Delta$ and the corresponding eigenvalues are $(n\pi)^2$. And indeed w_n form an orthogonal basis of $L^2((0,1))$ – and the expansion in that basis is the (sinusoidal) Fourier series.

For $n \geq 2$ such explicit computations are *very* difficult (already: there are many more open sets Ω in $2D$ than in $1D$ – although observe Theorem 21.34 (which since we look at inverses, describes the smallest eigenvalue of $-\Delta$).

There is a method of computing the eigenvalues variationally, which is the method of *Rayleigh Quotient*s, which is defined as

$$\frac{\|\nabla w\|^2_{L^2(\Omega)}}{\|w\|_{L^2(\Omega)}}.$$

Theorem 21.41. *Let $\Omega \subset \mathbb{R}^n$ be an open, bounded set with smooth boundary $\partial\Omega$.*

There exists a minimizier of the functional

$$E(w) := \frac{\|\nabla w\|^2_{L^2(\Omega)}}{\|w\|_{L^2(\Omega)}},$$

in the the set of functions

$$Y := \{w \in W^{1,2}_0(\Omega), \quad \|w\|_{L^2(\Omega)} \neq 0\}.$$

And any minimizer w satisfies

$$\Delta w = \lambda_1 w,$$

where λ_1 is the smallest eigenvalue of $-\Delta$.

Proof. Let w_k be a minimizing sequence in Y of E, i.e.

$$0 < \inf_{w \in Y} E(w) = \lim_{k\to\infty} E(w_k).$$

We may assume that

$$\sup_{k\in\mathbb{N}} E(w_k) \leq \Lambda,$$

for some huge Λ (simply take $\Lambda := E(\varphi)$ where φ is any choice of a nonzero $C^\infty_c(\Omega)$-function).

Otherwise replacing w_k by $\frac{w_k}{\|w_k\|_{L^2(\Omega)}}$ we may assume

$$\|w_k\|_{L^2(\Omega)} \equiv 1 \quad \forall k \in \mathbb{N}.$$

Then we have

$$\sup_k \|w_k\|_{L^2(\Omega)} + \|\nabla w_k\|_{L^2(\Omega)} \leq \Lambda,$$

and in particular w_k is uniformly bounded in $W^{1,2}_0(\Omega)$. By reflexivity there must exists a subsequence w_k converging weakly to $\bar{w}$ in $W^{1,2}_0(\Omega)$, and by Rellich's theorem strongly in $L^2(\Omega)$. We conclude that $\|\bar{w}\|_{L^2(\Omega)} = 1$. In particular $\bar{w} \in Y$. Moreover we have by weak lower semicontinuity of the norm

$$E(\bar{w}) = \|\nabla w\|^2_{L^2(\Omega)} \leq \liminf_{k\to\infty} \|\nabla w_k\|^2 = \liminf_{k\to\infty} E(w_k) = \inf_{w\in Y} E(w).$$

Since $w \in Y$ we have that $\bar{w}$ is indeed a minimizer of E.

Now let $\varphi \in C^\infty_c(\Omega)$ and consider

$$E(w + t\varphi) = \frac{\|\nabla w + t\nabla\varphi\|^2_{L^2(\Omega)}}{\|w + t\varphi\|_{L^2(\Omega)}}.$$

For small t we surely have $\|w+t\varphi\|_{L^2(\Omega)} \neq 0$ since $\|w\|_{L^2} = 1$, so this is all well-defined. By minimality we know

$$t \mapsto E(w+t\varphi)$$

has a minimum at $t=0$, so

$$0 = \frac{d}{dt}\Big|_{t=0} E(w+t\varphi) = \frac{2\int_\Omega \nabla w \cdot \nabla\varphi \|w\|^2_{L^2\Omega} - 2\|\nabla w\|^2_{L^2(\Omega)} \int w\varphi}{\|w\|_{L^2(\Omega)}} \overset{\|w\|_{L^2}=1}{=} 2\int_\Omega \nabla w \cdot \nabla\varphi - 2\|\nabla w\|^2_{L^2(\Omega)} \int w\varphi.$$

We conclude that

$$-\Delta w = \|\nabla w\|^2_{L^2(\Omega)} w.$$

So w is indeed an eigenfunction of the eigenvalue $\lambda := \|\nabla w\|^2_{L^2(\Omega)}$ (whatever that value is).

The theorem claims that λ is the *smallest* eigenvalue.

So assume $u \in W^{1,2}_0(\Omega) \setminus \{0\}$ satisfies

$$-\Delta u = \mu u \quad \text{in } \Omega$$

We can normalize to assume that $\|u\|_{L^2(\Omega)} = 1$.

Then we have

$$\lambda = \|\nabla w\|_{L^2(\Omega)} = E(w) \leq E(u) = \int_\Omega \nabla u \cdot \nabla u = \int_\Omega -\Delta u \, u = \mu \int_\Omega |u|^2 = \mu.$$

So we have shown $\lambda \leq \mu$ for any eigenvalue μ of Δ. □

Alright, so this is how we find the first eigenvalue, how do we find the next eigenvalue?

Let $w_1, w_2, \ldots, w_n$ be the first n eigenfunction we have found already (for now $n=1$). To find the $n+1$st Eigenfunction w_{n+1}, compute the infimum of the Raleigh quotient, only we minimize in the set

$$Y_{n+1} := \{w \in W^{1,2}_0(\Omega), w \in \{w_1, \ldots, w_n\}^\perp, \|w\|_{L^2(\Omega)} \neq 0\}.$$

Its now a nice but not completely trivial exercise to show this works:

Exercise 21.42. *Show that the minimum E in Y_{n+1} is attained, and that it computes the $n+1$-st eigenvalue.*

22. Semigroup theory

As references we refer to [Evans, 2010, §7.4] and [Cazenave and Haraux, 1998].

We could look at $(\partial_t - \Delta)\, u = 0$ and naively we should have

$$u = e^{t\Delta} u(0). \tag{22.1}$$

We can make this precise with the help of the Fourier Transform.

Is there a similar relation if we look at an elliptic operator L instead of Δ – or work with Δ on a domain where there is no Fourier transform? The first issue is that while we prefer to work in $L^2(\Omega)$, Δ is not defined everywhere.

22.1. **Unbounded Operators.** Let us introduce here the notion of *unbounded operators*.

Let X, Y be a Banach space. An operator

$$T\colon D(T) \subset X \to Y \tag{22.2}$$

is called linear, if and only if $D(T)$ is a linear subspace and T is linear on $D(T)$. We say T is *densely defined*, if

$$\overline{D(T)} = X. \tag{22.3}$$

T is *bounded* (or continuous), if and only if

$$\|T\| := \sup_{\|x\| \le 1} \|Tx\| < \infty. \tag{22.4}$$

Otherwise it is called *unbounded*.

examples

(1) $X = L^2(\mathbb{R}^n)$, $T = \Delta$, $D(T) = H^2(\mathbb{R}^n)$ or $D(T) = C^\infty$.
(2) $X = C^0([0,1])$, $D(T) = X$, $K \in C^0([0,1] \times [0,1])$

$$Tu(x) = \int_0^1 K(x,y)u(y)\, dy \tag{22.5}$$

is bounded.

We use the following notation.

$$G(T) = \{(u, Tu) \subset X \times X\colon u \in D(T)\} \tag{22.6}$$

is the *graph of T*,

$$R(T) = \{Tu\colon u \in D(T)\} \tag{22.7}$$

the *range of T*. An *extension of T* is

$$\tilde{T}\colon D(\tilde{T}) \subset X \to X, \tag{22.8}$$

such that

$$D(T) \subset D(\tilde{T}) \quad \text{and} \quad Tu = \tilde{T}u \quad \forall u \in D(T). \tag{22.9}$$

T is called *closed*, if $G(T)$ is closed in $X \times X$. That is, whenever $(u_k)_{k\in\mathbb{N}} \subset D(T)$ and $u, g \in X$ with $u_k \xrightarrow{k\to\infty} u$ in X, $Tu_k \xrightarrow{k\to\infty} g \in X$ we have $u \in D(T)$ and $g = Tu$.

T is called *closable*, if there exists a closed extension $\tilde{T}$. Recall the Closed Graph Theorem Theorem 20.23 – which shows that if $D(T) = X$ then closedness is the same as boundedness.

22.2. **semigroup setup.** Generally: Let X be a real Banach space and a linear map A,

$$A\colon D(A) \subset X \to X, \tag{22.10}$$

where $D(A)$ is the domain of A, a linear (usually dense) subset of X. We are looking for solutions $u \in C^1((0,T),X)$ of

$$\begin{aligned} \dot{u} &= Au, \quad t \in (0,T),\\ u(0) &= \varphi. \end{aligned} \tag{22.11}$$

A is in general not bounded, but closed. Assume there exists a solution to (22.11), then

$$T(t)\varphi := u(t) \tag{22.12}$$

defines an operator. Resonable properties of $T\colon$ are

- $T(t)\colon X \to X$ is linear.
- $T\colon [0,T) \to L(X)$. (hopefully)
- $T(0) = \mathrm{id}\,$,
- $T(t+s) = T(t) \circ T(s)$, (from uniqueness hopefully)
- $t \mapsto T(t)\varphi$ is continuous.

The latter three properties are characteristic for a *semigroup*.

Assume now that we have a semigroup

$$T\colon [0,\infty) \times X \to X. \tag{22.13}$$

Then we find some A such that T is the semigroup of A. A will then be called the *generator* of T.

Indeed, let $u(t) := T(t)\varphi$. Then,

$$\begin{aligned} \dot{u}(t) &= \lim_{s\to 0} \frac{u(t+s)-u(t)}{s} = \lim_{s\to 0} \frac{T(t+s)\varphi - T(t)\varphi}{s}\\ &= \lim_{s\to 0} \frac{T(s)-T(0)}{s} u(t)\\ &\equiv Au(t). \end{aligned} \tag{22.14}$$

Hence let

$$Au = \lim_{s\to 0} \frac{T(s)-T(0)}{s} u, \tag{22.15}$$

whenever the limit exists. Call $D(A)$ the set of $u \in X$ where this limit exists.

One might conjecture there is some sort of equivalence between generators A and semigroups T.

Questions: Which generators A allow semigroups? Which generators are obtained by semigroups?

The main theorem which gives us an answer to this question is the *Hille-Yoshida Theorem*, Theorem 22.20 and 22.22. For Schrödinger equations this is (a generalization of) *Stones' theorem*, it also appears under the name *Lumer–Phillips theorem*.

22.3. **m-dissipative operators.** Take an operator

$$A : D(A) \subset X \to X, \tag{22.16}$$

where X is a Banach space and $D(A)$ a linear subspace, e.g. $X = L^2$ and $D(A) = H^2$ and $A = \Delta$. The norm of our space X is the L^2-norm, and then A is not a bounded operator. On the other hand, since C_c^∞ is dense in L^2, $D(A)$ is dense in L^2 (everything with respect to the L^2-norm).

We want to solve (i.e. find $u(t)$) such that

$$\begin{cases} u'(t) = Au(t), \quad t > 0 \\ u(0) = \varphi \end{cases} \tag{22.17}$$

22.3.1. *The (easy) case of linear bounded operator A.* If A is simply a $\mathbb{R}^{n\times n}$-matrix, the situation is easy.

Example 22.1 (finite dimensional case)**.** Let $X = \mathbb{R}^n$ or $\mathbb{C}^n$, $A\colon X \to X$ linear (and thus bounded), then

$$u(t) = e^{tA}\varphi \tag{22.18}$$

is the unique solution to (22.17), where

$$e^{tA} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k. \tag{22.19}$$

Now let X be a general Banach space and $A \in L(X)$, where $L(X)$ is the space of bounded linear operators from X to X. I.e. A is *bounded*. Then e^{tA} still makes sense

Exercise 22.2. *Let $A, B \in L(X)$. Show*

(1) e^A converges absolutely, where

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

(2) We have the estimate

$$\|e^A\|_{L(X)} \leq e^{\|A\|_{L(X)}}.$$

(3) $e^0 = \mathrm{id}$,

(4) *if A and B commute, i.e. $AB = BA$, then*[48] $e^{A+B} = e^A e^B$
(5) $e^{-A} = \left(e^A\right)^{-1}$.
(6) *If we set* $f(t) := e^{tA}$ *then* $f \in C^\infty(\mathbb{R})$ *and we have* $f'(t) = e^{tA}A = Ae^{tA}$.

Theorem 22.3. *Let* $A \in L(X)$, $\varphi \in X$, $T > 0$. *Then there exists a unique solution* $u \in C^1((0,T), X)$ *of*

$$\begin{cases} u'(t) = Au(t) & t \in (0,T) \\ u(0) = \varphi. \end{cases}$$

Proof. Put

$$u(t) = e^{tA}\varphi. \tag{22.20}$$

Then

$$u'(t) = e^{tA}A\varphi = Au(t). \tag{22.21}$$

For a second solution v set

$$w(t) = e^{-tA}v(t), \tag{22.22}$$

then $w'(t) = 0$ and hence $w(t) = w(0) = \varphi$. □

Recall Section 22.1 for the notion of unbounded operators.

22.3.2. *Notion of m-dissipative operators.* Let $(X, \|\cdot\|)$ be a Banach space, and $D(A) \subset X$ a dense subspace.

Let $A\colon D(A) \subset X \to X$ be linear (but not necessarily bounded, nor closed).

Definition 22.4. A is *dissipative*, if

$$\|u - \lambda Au\| \geq \|u\| \quad \forall u \in D(A), \lambda > 0. \tag{22.23}$$

A is called *accretive*, if $-A$ is dissipative.

Lemma 22.5. *Let X be a Hilbert*[49] *space,*

$$A\colon D(A) \subset X \to X \tag{22.24}$$

linear, then A is dissipative if and only if

$$\operatorname{Re}\langle u, Au\rangle \leq 0 \quad \forall u \in D(A). \tag{22.25}$$

[48] If $AB \neq BA$ there is a formula in terms of *commutators* $[A,B] = AB - BA$, the *Baker–Campbell–Hausdorff formula*

[49] i.e. there exists a scalar product $\langle , \rangle : X \times X \to \mathbb{C}$ such that $\|v\|^2 = \langle v, v\rangle$. For $L^2(\Omega)$ the scalar product $\langle f, g\rangle := \int_\Omega fg$ (real) or $\langle f, g\rangle := \int_\Omega f\overline{g}$ (complex)

Proof of Lemma 22.5. Assume A dissipative, then:

$$\|u\|^2 + \lambda^2\|Au\|^2 - 2\lambda \mathrm{Re}\,\langle u, Au\rangle - \|u\|^2 = \|u - \lambda Au\|^2 - \|u\|^2 \geq 0. \tag{22.26}$$

Dividing by λ and letting $\lambda \to 0$ gives

$$\mathrm{Re}\,\langle u, Au\rangle \leq 0. \tag{22.27}$$

For the converse, assume that

$$\mathrm{Re}\,\langle Au, u\rangle \leq 0, \tag{22.28}$$

then

$$\|u - \lambda Au\|^2 = \|u\|^2 + \lambda^2\|Au\|^2 - 2\lambda \mathrm{Re}\,\langle u, Au\rangle \geq \|u\|^2. \tag{22.29}$$

□

Example 22.6.
- *Heat equation* $(\partial_t - \Delta)u = 0$:
  Then $A = \Delta$, $X = L^2(\mathbb{R}^n)$, $D(A) = H^2(\mathbb{R}^n)$, then
  $$\langle u, \Delta u\rangle = -\int_{\mathbb{R}^n} |\nabla u|^2 \leq 0.$$
- *Schrödinger equation* $(\partial_t - i\Delta)u = 0$. $A = i\Delta$, $D(A) = H^2(\mathbb{R}^n)$,
  $$\langle u, \pm i\Delta u\rangle = \mp i \int_{\mathbb{R}^n} |\nabla u|^2$$
  and hence the real part is 0. That is, both $i\Delta$ and $-i\Delta$ are dissipative.

Definition 22.7 (m-dissipative)**.** A linear operator $A\colon D(A) \subset X \to X$ is called *m-dissipative*, if A is dissipative and $I - \lambda A$ is surjective for all $\lambda > 0$.

We call A *m-accretive*, if $-A$ is m-dissipative.

Our later goal is to show that for any m-dissipative A we can define (some sort of) e^A., which we will do for Hille-Yoshida Theorem 22.20.

Observe that for the notion of m-dissipative the right *choice of $D(A)$* becomes relevant. E.g., if we choose $D(\Delta) = C_c^\infty(\Omega)$ then Δ is dissipative, but not m-dissipative; but if we choose $D(\Delta) = H^2(\Omega)$ then Δ is m-dissipative, see Example 22.11.

Exercise 22.8. *Show that if $A : D(A) \subset X \to X$ is m-dissipative, then $I - \lambda A : D(A) \to X$ is bijective and has a continuous inverse $I - \lambda A : X \to D(A)$.*

Set

$$J_\lambda = (I - \lambda A)^{-1}\colon X \to D(A). \tag{22.30}$$

Then (22.23) implies for m-dissipative A

$$\|J_\lambda v\| \leq \|v\| \quad \forall v \in X. \tag{22.31}$$

Lemma 22.9. *Let A be dissipative, then A is m-dissipative if and only if there exists $\lambda_0 > 0$ such that $I - \lambda_0 A$ is surjective.*

Proof. Let $\lambda \in (0, \infty)$ and $v \in X$. Our goal is to find $u \in D(A)$ such that $u - \lambda A u = v$. Observe that it is equivalent to find u such that

$$u - \lambda_0 A u = \frac{\lambda_0}{\lambda} v + \left(1 - \frac{\lambda_0}{\lambda}\right) u, \tag{22.32}$$

or equivalently (recall $J_{\lambda_0} = (I - \lambda_0 A)^{-1} : X \to D(A)$)

$$u = J_{\lambda_0}\left(\frac{\lambda_0}{\lambda} v + \left(1 - \frac{\lambda_0}{\lambda}\right) u\right) =: F(u). \tag{22.33}$$

That is, the desired u is a fixed point of $u = F(u)$.

First we show that $F : X \to X$ is a contraction for $\lambda > \frac{\lambda_0}{2}$. (Recall, that we need to find $u \in D(A)$).

Observe, by (22.31),

$$\|F(u) - F(w)\| = \left\|J_{\lambda_0}\left(\left(1 - \frac{\lambda_0}{\lambda}\right)(u - w)\right)\right\| \leq \left|1 - \frac{\lambda_0}{\lambda}\right| \|u - w\|. \tag{22.34}$$

Hence F is a contraction, whenever $\lambda > \lambda_0/2$.

Then we apply Banach Fixed Point theorem , on X (*not* $D(A)$ which is not closed in general).

Let $u \in X$ be the fixed point, then we have $u = F(u) \in J_{\lambda_0}(X) \subset D(A)$.

That is, there is a unique $u \in D(A)$ with $F(u) = u$. This $I - \lambda A$ is surjective, whenever $\lambda > \frac{\lambda_0}{2}$.

Iterating this argument, e.g setting $\lambda_{i+1} := \frac{2}{3}\lambda_i$, we find that for any $\lambda > \lambda_i$, $i \in \mathbb{N}$, $I - \lambda A$ is surjective, and letting $i \to \infty$ we see that for any $\lambda > 0$ we have $I - \lambda A$ is surjective. □

Proposition 22.10. *All m-dissipative operators $(A, D(A))$ are closed.*

Proof. Let $u_k \to u$ in X and $A u_k \to g$ in X. We need to show that $u \in D(A)$ and $Au = g$.

Observe that $(I - A)u_k \to u + g$. Since A is m-dissipative, $J_1 = (I - A)^{-1}$ exists and is a continuous map $X \to D(A)$. Thus

$$u \xleftarrow{k\to\infty} u_k = J_1((I - A)u_k) \xrightarrow{k\to\infty} J_1(u - g) = J_1 u - J_1 g \in D(A)$$

That is, $u = J_1 u - J_1 g \in D(A)$, and applying $I - A$ we obtain $(I - A)u = u - g$, i.e. $Au = g$. □

Example 22.11. Let Ω be a bounded set, $\partial\Omega \in C^\infty(\Omega)$, or $\Omega = \mathbb{R}^n$.

Set $X = L^2(\Omega)$, $A = \Delta$, $D(A) = H^1_0(\Omega) \cap H^2(\Omega)$. Then A is m-dissipative.

Proof. We already know that A is dissipative (Example 22.6), so by Lemma 22.9 we only need to show that

$$\forall v \in L^2(\Omega)\ \exists u \in H^2(\Omega) \cap H^1_0(\Omega)\colon u - \Delta u = v. \tag{22.35}$$

There are several ways to do this. Observe that -1 is not an eiganvalue of $-\Delta$ (since $-\Delta$ is positive, Lemma 21.38, so we could use Fredholm alternative Theorem 21.13 to find $u \in H^1_0(\Omega)$.

Another option is to go variational, computing via the direct method of Calculus of Variations, Theorem 12.30, that the minimizer $u \in H^1_0(\Omega)$ of

$$\mathcal{E}(u) := \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{2}\int u^2 - \int uv$$

exists.

In both cases we find $u \in H^1_0(\Omega)$ that solve the equation. In particular we have

$$\Delta u \in L^2(\Omega).$$

From L^2-regularity theory in PDE we then conclude $u \in H^2(\Omega)$. $\square$

Proposition 22.12. *Let A be m-dissipative, then*

$$\forall u \in \overline{D(A)}\colon \quad \|J_\lambda u - u\| \xrightarrow{\lambda \to 0} 0. \tag{22.36}$$

Proof. Observe that by (22.31), $J_\lambda - I : X \to X$ is bounded linear operator, i.e.

$$\|J_\lambda - I\| \le \|J_\lambda\| + \|I\| \le 2. \tag{22.37}$$

By density $D(A) \subset \overline{D(A)}$, it thus suffices to prove the result for $u \in D(A)$. Since $J_\lambda = (I - \lambda A)^{-1}$ using again (22.31)

$$\|J_\lambda u - u\| = \|J_\lambda \left(u - (I - \lambda A)u\right)\| \le \lambda\|Au\| \to 0, \quad \lambda \to 0. \tag{22.38}$$

$\square$

Observe that from (22.31) we have the following useful observations

$$J_\lambda - I = J_\lambda(I - (I - \lambda A)) = \lambda J_\lambda A,$$

and

$$J_\lambda - I = \left((I - (I - \lambda A))\right) J_\lambda = \lambda A J_\lambda.$$

Set

$$A_\lambda := AJ_\lambda = \frac{1}{\lambda}(J_\lambda - I). \tag{22.39}$$

This $A_\lambda \in L(X)$ will serve as an "approximation" for A, so that we can make (certain) sense of an operator e^{tA} in terms of $\lim_{\lambda\to 0} e^{tA_\lambda}$. This is justified by the following

Proposition 22.13. *Let A be m-dissipative and $\overline{D(A)} = X$. Then*

$$A_\lambda u \xrightarrow{\lambda\to 0} Au, \quad \forall u \in D(A). \tag{22.40}$$

Proof. We have

$$(I - \lambda A)A = A(I - \lambda A). \tag{22.41}$$

Thus, multiplying both sides with J_λ from the left and also from the right, we have $A_\lambda = AJ_\lambda = J_\lambda A$.

Now observe that by Proposition 22.12,

$$J_\lambda Au \xrightarrow{\lambda\to 0} Au, \tag{22.42}$$

since $D(A)$ is dense in X □

22.4. **Semigroup Theory.** Let X be a Banach space. A *semigroup* is an operator

$$T\colon [0,\infty) \to L(X), \tag{22.43}$$

such that

(i) $T(0) = I$,
(ii) $T(t+s) = T(t)T(s)$.

T is called *C^0-semigroup* (*strongly continuous semigroup*), if

(iii) $\lim_{t\to 0} \|T(t)u - u\|_X = 0 \quad \forall u \in X$.

Note, that by (ii) we necessarily have $T(s)T(t) = T(t)T(s)$.

Example 22.14. (1) $A \in L(X)$, $T(t) = e^{tA}$.
(2) $X = L^p(\mathbb{R})$, $p \in [1,\infty]$.

$$T(t)u(x) = u(t+x). \tag{22.44}$$

If $p < \infty$, then T is a continuous semigroup, since C_c^∞ is dense and hence for $u \in L^p$ and $\varepsilon > 0$ there exists $f \in C_c^\infty$ with

$$\|f - u\|_p < \varepsilon/3. \tag{22.45}$$

We have for all small t,

$$\sup_x |f(x-t) - f(x)| < t\|\nabla f\|_\infty < \varepsilon/3, \tag{22.46}$$

and thus for $t \ll 1$

$$\left(\int_{\mathbb{R}} |T(t)f - f|^p\right)^{\frac{1}{p}} < \frac{\varepsilon}{3}\left(\operatorname{diam}(\operatorname{supp} f) + 1\right). \tag{22.47}$$

Moreover, by the definition of $T(t)$

$$\|T(t)(u-f)\|_{p,\mathbb{R}^n} = \|u-f\|_{p,\mathbb{R}^n} \leq \frac{\varepsilon}{3}.$$

Thus,

$$\begin{aligned}\|T(t)u-u\|_p &\leq \|T(t)f-f\|_p + \|T(t)(u-f)\|_p + \|u-f\|_p\\ &\leq \varepsilon\end{aligned}$$

The situation is different for $p=\infty$: let $u=\chi_{[0,1]}$, then

$$\|u-T(t)u\|_\infty = \sup_x |u(x)-u(x+t)| \geq 1 \quad \forall t>0. \tag{22.48}$$

Thus T is no C^0-semigroup for $p=\infty$.

Proposition 22.15. *Let $T(t)$ be a C^0-semigroup on X. Then $\exists M \geq 1$ and $\omega \in \mathbb{R}$ such that*

$$\|T(t)\|_{L(X)} \leq Me^{\omega t} \quad \forall t>0. \tag{22.49}$$

Proof. We show that there exists $\delta>0$ such that

$$M := \sup_{0<t<\delta} \|T(t)\| < \infty. \tag{22.50}$$

If this was not the case, then there exists a sequence $t_n \to 0$ with

$$\|T(t_n)\| \to \infty. \tag{22.51}$$

Recall Banach-Steinhaus Theorem, Theorem 12.18: If for a sequence $A_n \in L(X)$ we have

$$\forall u \in X\colon \sup_n \|A_n u\| < \infty, \tag{22.52}$$

then $\sup_n \|A_n\| < \infty$.

Hence, if (22.51) holds, then there must be $u \in X$ such that $\|T(t_n)u\| \to \infty$. But this is in contradiction to the C^0-property. Hence (22.51) cannot hold, i.e. (22.50) must be true.

Now let $t>0$, then there exists $n \in \mathbb{N}$ and $s \in (0,\delta)$, such that

$$t = n\delta + s. \tag{22.53}$$

Then by the semigroup property

$$T(t) = \underbrace{T(\delta)\circ\cdots\circ T(\delta)}_{\text{n times}} \circ T(s). \tag{22.54}$$

That is, with (22.50),

$$\|T(t)\| \leq \|T(\delta)\|^n \|T(s)\| \leq M^{n+1} \leq MM^{\frac{t}{\delta}} = Me^{t\log\frac{M}{\delta}}. \tag{22.55}$$

□

Proposition 22.16. *Let $T(t)$ be a C^0-semigroup. Then the map*

$$(t, u) \mapsto T(t)u \tag{22.56}$$

is continuous.

Exercise 22.17. *Prove Proposition 22.16.*

Definition 22.18. Let $T(t)$ be a C^0-semigroup. Then

$$\omega_0 = \inf\{w \in \mathbb{R} \colon \exists M \geq 1, \|T(t)\| \leq Me^{\omega t}\} \tag{22.57}$$

ist called the *growth bound* of the semigroup.

Definition 22.19. A C^0-semigroup is called *contraction semigroup*, if

$$\forall t > 0 \colon \|T(t)\| \leq 1. \tag{22.58}$$

Recall that

$$\|J_\lambda\| \leq 1, \quad \|A_\lambda\| \leq \frac{2}{\lambda}, \tag{22.59}$$

where

$$A_\lambda := AJ_\lambda = \frac{1}{\lambda}(J_\lambda - I). \tag{22.60}$$

We define

$$T_\lambda(t) = e^{tA_\lambda}. \tag{22.61}$$

For any $\lambda > 0$, T_λ is a C^0-semigroup (because $A_\lambda \in L(X)$). Moreover, for any $\lambda > 0$ we have that T_λ it is a contraction semigroup:

$$\|T_\lambda(t)\|_{L(X)} = \|e^{tJ_\lambda \frac{1}{\lambda}} e^{-\frac{t}{\lambda} I}\| = e^{-\frac{t}{\lambda}} \|e^{\frac{t}{\lambda} J_\lambda}\| \leq e^{-\frac{t}{\lambda}} e^{\frac{t}{\lambda}} = 1. \tag{22.62}$$

Observe that in contrast, $e^{t\|A_\lambda\|}$ is in generally not uniformly bounded w.r.t $\lambda \to 0$.

The semigroup for an m-dissipative operator A will be constructed out of $T_\lambda \xrightarrow{\lambda \to 0} T(t)$. The following is the (first part) of the main theorem of the semigroup theory:

Theorem 22.20 (Hille Yoshida (Part I))**.** *Let $A \colon D(A) \subset X \to X$ m-dissipative and densely defined. Then for all $u \in X$ the limit*

$$T(t)u = \lim_{\lambda \to 0} T_\lambda(t)u \tag{22.63}$$

exists and the convergence is uniform (w.r.t. t) on time-intervals of the form $[0, T]$.

Furthermore $(T(t))_{t \geq 0}$ is a contraction semigroup, $T(t)(D(A)) \subset D(A)$, and for all $u \in D(A)$,

$$u(t) := T(t)u \tag{22.64}$$

is the unique solution $u \in C^0([0,\infty), D(A)) \cap C^1((0,\infty), X)$ *to*

$$\begin{cases} \dot{u}(t) & = Au(t) \quad t > 0 \\ u(0) & = u \end{cases} \tag{22.65}$$

Proof. **Step (1): On the contraction semigroup property**

First we show that for $\mu, \lambda, t > 0$

$$\|T_\lambda(t)u - T_\mu(t)u\|_X \leq t\,\|A_\mu u - A_\lambda u\|_X. \tag{22.66}$$

Indeed, observe that (A_λ and A_μ commute!)

$$\begin{aligned} T_\lambda(t) - T_\mu(t) =& e^{tA_\lambda} - e^{sA_\mu} = \int_0^t \frac{d}{ds}\left(e^{(t-s)A_\mu} e^{sA_\lambda}\right)\, ds \\ =& \int_0^t e^{(t-s)A_\mu} e^{sA_\lambda}\, ds\ (A_\lambda - A_\mu) \end{aligned}$$

Consequently, (recall (22.62))

$$\begin{aligned} \|T_\lambda(t)u - T_\mu(t)u\|_X \leq& \int_0^t \underbrace{\|e^{(t-s)A_\mu}\|_{L(X)}}_{\leq 1} \underbrace{\|e^{sA_\lambda}\|_{L(X)}}_{\leq 1}\, ds\, \|(A_\lambda - A_\mu)u\|_X \\ \leq& t\|(A_\lambda - A_\mu)u\|_X. \end{aligned}$$

Thus (22.66) is established.

From (22.66) we conclude for any fixed $u \in D(A)$, using also Proposition 22.13, that $(T_\lambda(t)u)_{\lambda>0}$ is a Cauchy-sequence as $\lambda \to 0$ in X w.r.t λ, and this Cauchy sequence is uniform in $t \in [0,T]$ for any fixed $T < \infty$.

Hence the proposed limit $T(t)u := \lim_{\lambda\to 0} T_\lambda(t)u$ exists for any $u \in D(A)$, and Precisely,

$$\|T_\lambda(t)u - T(t)u\|_X \leq t\|(A_\lambda - A)u\|_X \quad \forall u \in D(A).$$

Clearly, $T(t)$ a linear operator, and from the above estimate together with (22.62) we find

$$\begin{aligned} \|T(t)u\|_X \leq& \|T_\lambda(t)u - T(t)u\|_X + \|T_\lambda(t)u\|_X \\ \leq& \|T_\lambda(t)u - T(t)u\|_X + \|u\|_X \\ \leq& t\|(A_\lambda - A)u\|_X + \|u\|_X. \end{aligned}$$

Letting $\lambda \to 0$, we find

$$\|T(t)u\|_X \leq \|u\|_X \quad \forall u \in D(A).$$

That is $T(t) : D(A) \to X$ is a linear bounded operator; since $D(A)$ is a dense set in X, $T(t)$ can be extended to a linear bounded operator on all of X.

Now let $u \in X$ with approximating sequence $u_n \in D(A)$.

$$\begin{aligned}\|T_\lambda(t)u - T(t)u\| \leq& \|T_\lambda(t)u - T_\lambda(t)u_n\| + \|T_\lambda(t)u_n - T(t)u_n\| \\ &+ \|T(t)(u_n - u)\| \\ \leq& 2\|u_n - u\| + \|T_\lambda(t)u_n - T(t)u_n\| \\ \leq& 2\|u_n - u\| + t\|(A_\lambda - A)u_n\|_X\end{aligned} \tag{22.67}$$

Hence by (22.66) we find $T_\lambda(t)u \to T(t)u$ for all $u \in X$, and indeed the above estimate implies even uniformity in t: namely, for any $t_0 > 0$,

$$\sup_{t\in[0,t_0)} \|T_\lambda(t)u - T(t)u\| \xrightarrow{\lambda\to 0} 0 \quad \forall u \in X. \tag{22.68}$$

Furthermore we have the semigroup property

$$\begin{aligned}\|T(t)T(s)u - T(t+s)u\| \leq& \|T(t)T(s)u - T(t)T_\lambda(s)u\| \\ &+ \|T(t)T_\lambda(s)u - T_\lambda(t)T_\lambda(s)u\| \\ &+ \|T_\lambda(t+s)u - T(t+s)u\| \\ \to& 0.\end{aligned} \tag{22.69}$$

As for the C^0-continuity, we have

$$\|T(t)u - u\|_X \leq \|T_\lambda(t)u - u\|_X + \|T_\lambda(t)u - T(t)u\|_X.$$

By (22.68), for any $\varepsilon > 0$ and any $t_0 > 0$ there exists $\lambda > 0$ such that

$$\sup_{t\in[0,t_0)} \|T_\lambda(t)u - T(t)u\|_X < \frac{1}{2}\varepsilon.$$

On the other hand for this fixed $\lambda > 0$ there exists $t_1 \in (0, t_0)$ such that

$$\sup_{t<t_1} \|T_\lambda(t)u - u\|_X < \frac{\varepsilon}{2}.$$

In particular,

$$\sup_{t<t_1} \|T(t)u - u\|_X \leq \varepsilon.$$

That is $T(t)$ is a contractive C^0-semigroup.

Next we show that $\underline{T(t) \text{ maps } D(A) \text{ to } D(A)}$. First we observe that T_λ maps $D(A)$ to $D(A)$. Indeed, since we can write $e^{J_\lambda} = I + J_\lambda B$ (with B convergent since $J_\lambda \in L(X)$) and since J_λ maps X into $D(A)$ we have that e^{J_λ} maps $D(A)$ into $D(A)$. Moreover, by (22.60) we have $T_\lambda(t) = e^{tA_\lambda} = e^{t\frac{1}{\lambda}(J_\lambda - I)} = e^{-\frac{t}{\lambda}}e^{J_\lambda}$, so $T_\lambda(t)$ maps $D(A)$ into $D(A)$. Now let $u \in D(A)$, then $T_\lambda(t)u \xrightarrow{\lambda\to 0} T(t)u$, and $AT_\lambda(t)u = T_\lambda(t)Au \to T(t)Au$. By Proposition 22.10, A is a closed operator, which implies that $T(t)u \in D(A)$ and $\lim_{\lambda\to 0} AT_\lambda(t)u = AT(t)u$, and in particular,

$$AT(t)u = T(t)Au \quad \forall u \in D(A). \tag{22.70}$$

Step (2): On the equation (22.65)

Let $u \in D(A)$ and set

$$u_\lambda(t) = e^{tA_\lambda}u. \tag{22.71}$$

Then

$$\frac{d}{dt}u_\lambda = e^{tA_\lambda}A_\lambda u = T_\lambda(t)A_\lambda u. \tag{22.72}$$

Equivalently, for $u \in D(A)$ using $A_\lambda u \to Au$ and $T_\lambda \to T$,

$$u(t) \leftarrow u_\lambda(t) = u + \int_0^t T_\lambda(s)A_\lambda u\ ds \to u + \int_0^t T(s)\ Au\ ds. \tag{22.73}$$

Thus $u(\cdot) \in C^1$ and (cf. (22.70))

$$\dot{u}(t) = T(t)Au = Au(t). \tag{22.74}$$

Uniqueness proceeds as in Theorem 22.3. □

22.4.1. *Generators of semigroups.* Let $T(t)$ be a contraction semigroup. Define

$$D(A) := \left\{u \in X\colon \lim_{h\to 0^+} \frac{T(h)u - u}{h} \text{ exists}\right\}. \tag{22.75}$$

For $u \in D(A)$ set

$$Au = \lim_{h\to 0^+} \frac{T(h)u - u}{h}. \tag{22.76}$$

Example 22.21. $X = C_{ub}(\mathbb{R})$ be the set of uniformly continuous, bounded functions with the A^∞-norm.

$$T(t)u(x) := u(x+t). \tag{22.77}$$

Then $T(t)$ is a contraction semigroup. Then

$$Au = u', \quad D(A) = \{u, u' \in C_{ub}(\mathbb{R})\}. \tag{22.78}$$

Proof. It is clear that $u, u' \in C_{ub}(\mathbb{R})$ implies

$$\left\| \frac{u(x+h) - u(x)}{h} - u'(x)\right\|_\infty \to 0. \tag{22.79}$$

Now let $u \in D(A)$, then $u'_+ \in C_{ub}(\mathbb{R})$ and hence $u'_+ = u' \in C_{ub}(\mathbb{R})$. □

Theorem 22.22 (Hille Yoshida Part II)**.** *Let $T(t)$ be a contraction semigroup with generator A. Then A is m-dissipative and densely defined.*

Proof. (i) A is dissipative, i.e. for all $\lambda > 0$, $\|u - \lambda Au\| \geq 0$. Indeed,

$$\begin{aligned}\left\|u-\lambda\frac{T(h)u-u}{h}\right\| &\geq \left\|\left(1+\frac{\lambda}{h}\right)u\right\| - \left\|\frac{\lambda}{h}T(h)u\right\| \\ &= \left(1+\frac{\lambda}{h}\right)\|u\| - \frac{\lambda}{h}\|T(h)u\| \\ &\geq \left(1+\frac{\lambda}{h}\|u\| - \frac{\lambda}{h}\|u\|\right) = \|u\|.\end{aligned} \tag{22.80}$$

In the last step we used that $T(h)$ is contracting, i.e. $\|T(h)u\|_X \leq \|u\|_X$. Letting $h \to 0$ on the left hand side shows A is dissipative.

(ii) A is m-dissipative. It suffices to show that $(I - A)$ is surjective. Thus we want to find Ju, such that

$$(I-A)Ju = u. \tag{22.81}$$

Ansatz:

$$Ju = \int_0^\infty e^{-t}T(t)\ dt. \tag{22.82}$$

Why? Because (formally!) we know

$$(I-A)T(t)u = T(t)u - \partial_t(T(t)u)$$

This is equivalent to

$$(I-A)e^{-t}T(t)u = -\partial_t(e^{-t}T(t)u)$$

Integrating on both sides on $t \in (0,\infty)$ we then should get

$$(I-A)Ju = T(0)u = u.$$

To make this more precise, first observe

$$\|Ju\| \leq \int_0^\infty e^{-t}\|T(t)u\|\ dt \leq \|u\| \tag{22.83}$$

and hence $\|J\| \leq 1$. Now let us compute (with the semigroup property $T(h)T(t) = T(h+t)$ for $h, t > 0$),

$$\begin{aligned}(T(h)-I)\,Ju &= \int_0^\infty e^{-t}T(t+h)u\ dt - \int_0^\infty e^{-t}T(t)u\ dt \\ &= \int_h^\infty e^{-t+h}T(t)u\ dt - \int_0^\infty e^{-t}T(t)u\ dt \\ &= \int_0^\infty \left(e^{-t+h} - e^{-t}\right)T(t)u - \int_0^h e^{-t+h}T(t)u\ dt \\ &= (e^h-1)\int_0^\infty e^{-t}T(t)u\ dt - e^h\int_0^h e^{-t}T(t)u\ dt \\ &= (e^h-1)Ju - e^h\int_0^h e^{-t}T(t)u\ dt.\end{aligned} \tag{22.84}$$

Hence

$$\frac{T(h)-I}{h}Ju = \frac{e^h-1}{h}Ju - \frac{e^h}{h}\int_0^h e^{-t}T(t)u\,dt. \tag{22.85}$$

The right-hand side converges as $h \to 0$ (we use that T is continuous at 0), consequently so does the left-hand side. Thus $Ju \in D(A)$ and

$$AJu = Ju - u, \tag{22.86}$$

which is the claim.

(iii) $D(A)$ is dense. Let $u \in X$, then we claim that the following $u_h \in D(A)$ and $u_h \xrightarrow{h\to 0} u$ in X. Namely,

$$u_h := \frac{1}{h}\int_0^h T(s)u\,ds. \tag{22.87}$$

First we show that $u_h \xrightarrow{h\to 0} u$ holds. Indeed, we use again that T is continuous at 0,

$$\begin{aligned} \|u_h - u\| &= \left\|\frac{1}{h}\int_0^h (T(s)-I)\,u\,ds\right\| \\ &\leq \frac{1}{h}\int_0^h \|(T(s)-I)\,u\| \xrightarrow{h\to 0} 0. \end{aligned} \tag{22.88}$$

Now show that $u_h \in D(A)$ for all $h > 0$. Let $t << h$. We calculate

$$\begin{aligned} \frac{T(t)-I}{t}u_h &= \frac{1}{ht}\int_0^h T(t)\circ T(s)u\,ds - \frac{1}{ht}\int_0^h T(s)u\,ds \\ &= \frac{1}{ht}\int_t^{t+h} T(s)u\,ds - \frac{1}{ht}\int_0^h T(s)u\,ds \\ &= \frac{1}{ht}\int_h^{t+h} T(s)u\,ds + \frac{1}{ht}\int_t^h T(s)u\,ds \\ &\quad - \frac{1}{ht}\int_0^t T(s)u\,ds - \frac{1}{ht}\int_t^h T(s)u\,ds \\ &\xrightarrow{t\to 0} \frac{1}{h}T(h)u - \frac{1}{h}T(0)u \in X. \end{aligned} \tag{22.89}$$

Hence the left hand side converges in X, that is $u_h \in D(A)$. □

22.5. An example application of Hille-Yoshida.

Example 22.23. Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary $\partial\Omega$. We want to find solutions to the equation

$$\begin{cases} (\partial_t - L)u = 0 & \text{in } \Omega \times (0,\infty) \\ u = 0 & \text{in } (\partial\Omega) \times (0,\infty) \\ u = u_0 & \text{in } \Omega \times \{0\} \end{cases} \tag{22.90}$$

where the operator L given as

$$Lu(x) := \partial_i(a_{ij}(x)\partial_j u) + b_i(x)\partial_i u(x) + c(x)u(x)$$

has smooth (*time-independent*) coefficients $a, b, c \in C^\infty(\overline{\Omega})$, and is uniformly elliptic, i.e. for some $\lambda > 0$

$$a_{ij}(x)\xi_i\,\xi_j \geq \lambda|\xi|^2 \quad \forall x \in \Omega, \xi \in \mathbb{R}^n.$$

We claim that there exists a solution to (22.90) (in the sense of a semigroup) such that

$$\|u(t)\|_{L^2(\Omega)} \leq e^{ct}\|u_0\|_{L^2(\Omega)} \quad \forall t \in (1, \infty).$$

Proof. To see this let $X = L^2(\Omega)$, $D(A) := H^2(\Omega) \cap H^1_0(\Omega)$, and $A := L - \omega$ for an $\omega \geq 0$ yet to be chosen. Observe that if we show that A is m-dissipative then by Hille-Yoshida, Theorem 22.20, we find a contraction semigroup $T(t)$ such that for $\tilde{u}(t) := T(t)u_0$

$$\partial_t(\tilde{u}(t)) = A\tilde{u}(t)$$

with the contractive property

$$\|\tilde{u}(t)\|_{L^2(\Omega)} \leq \|u_0\|_{L^2(\Omega)}.$$

Setting $u(t) := e^{\omega t}\tilde{u}(t)$ then $u(0) = \tilde{u}(0) = u_0$ and the requested equation is satisfied

$$\partial_t(u(t)) = \omega e^{\omega t}\tilde{u}(t) + Ae^{\omega t}\tilde{u}(t) = Lu(t).$$

Moreover, we obtain

$$\|u(t)\|_{L^2(\Omega)} \leq e^{\omega t}\|u_0\|_{L^2(\Omega)}.$$

That is, we need to show that A is m-dissipative (clearly $D(A)$ is dense in L^2). □

First we would like to show that A is dissipative.

Lemma 22.24. *There exist $\omega_0 > 0$ such that for any $\omega \geq \omega_0$, A as above is dissipative.*

Proof. In view of Lemma 22.5 we need to show that

$$\langle Au, u\rangle_{L^2(\Omega)} \leq 0 \quad \forall u \in D(A).$$

Here is where we need to make use of ω:

$$\langle Au, u\rangle_{L^2(\Omega)} \leq -\int_{\mathbb{R}^n} a_{ij}\partial_i u\partial_j u + C(b)\left(\|\nabla u\|_{L^2}\|u\|_{L^2} + \|u\|^2_{L^2}\right) - \omega\|u\|_{L^2(\Omega)}$$

By ellipticity,

$$-\int_{\mathbb{R}^n} a_{ij}\partial_i u\partial_j u \leq -\lambda\|\nabla u\|_{L^2(\Omega)}$$

Moroever by Young's inequality, for any $\varepsilon > 0$,

$$\|\nabla u\|_{L^2}\|u\|_{L^2} \leq \varepsilon\|\nabla u\|^2_{L^2(\Omega)} + C\frac{1}{\varepsilon}\|u\|^2_{L^2(\Omega)},$$

so that for small enough $\varepsilon > 0$ we arrive at

$$\langle Au, u\rangle_{L^2(\Omega)} \leq -\frac{\lambda}{2}\|\nabla u\|^2_{L^2} + C\|u\|^2_{L^2} - \omega\|u\|_{L^2(\Omega)}$$

Choosing $\omega := C$ we conclude

$$\langle Au, u\rangle_{L^2(\Omega)} \leq -\frac{\lambda}{2}\|\nabla u\|_{L^2}^2 \leq 0. \tag{22.91}$$

that is, by Lemma 22.5, A is dissipative. □

The next step is to prove that A is actually m-dissipative.

In view of Lemma 22.9 it suffices to show that $I - \lambda_0 A$ is surjective for some $\lambda_0 > 0$.

That is, for some fixed $\lambda_0 > 0$ and $\omega > \omega_0$ we need to show that for any $f \in L^2(\Omega)$ there exists $u \in H^1_0(\Omega) \cap H^2(\Omega)$ with $(I - \lambda_0 A)u = f$, that is we need to find u such that

$$\partial_i(a_{ij}\partial_j u) = -\frac{1}{\lambda_0}f + (\frac{1}{\lambda_0} + \omega)u - b_j\partial_j u - cu \tag{22.92}$$

The $\partial_i(a_{ij}\partial_j u)$ is the leading order term, the remaining terms are lower order terms that can be dealt with by a fixed point argument. First we treat the leading order term:

Lemma 22.25. *Let $g \in L^2(\Omega)$*[50] *then there exists exactly one $u \in H^1_0(\Omega)$ with*

$$\partial_i(a_{ij}\partial_j u) = g \tag{22.93}$$

in distributional sense. Moreover,

$$\|u\|_{H^1(\Omega)} \leq C\,\|g\|_{L^2(\Omega)} \tag{22.94}$$

Proof. This can be done, e.g. by a variational argument, the direct method.

Let $\mathcal{E} : H^1_0(\Omega) \to \mathbb{R}$ be an energy defined as

$$\mathcal{E}(u) := \frac{1}{2}\int_\Omega a_{ij}\,\partial_i u\,\partial_j u + \int_\Omega g\,u.$$

- Assume that $\mathcal{E}(\cdot)$ has a minimizer $u \in H^1_0(\Omega)$, i.e. $\mathcal{E}(u) \leq \mathcal{E}(v)$ for all $v \in H^1_0(\Omega)$. Then $\mathcal{E}(u) \leq \mathcal{E}(u + t\varphi)$ for any $\varphi \in C_c^\infty(\Omega)$, $t \in \mathbb{R}$. Thus $t \mapsto \mathcal{E}(u + t\varphi)$ as a minimum at $t = 0$, and consequently, by Fermat's theorem (a_{ij} is symmetric!)
$$0 = \frac{d}{dt}\Big|_{t=0}\mathcal{E}(u + t\varphi) = \int_\Omega a_{ij}\,\partial_i u\,\partial_j\varphi + \int_\Omega g\,\varphi.$$
This holds for any $\varphi \in C_c^\infty(\Omega)$; That is, (22.93) is satisfied (in distributional sense). Indeed, we say that (22.93) is the *Euler-Lagrange equation* of the energy $\mathcal{E}$.
- $\mathcal{E}(\cdot)$ is coercive (it controls $\|\nabla u\|_{L^2(\Omega)}$.
$$\mathcal{E}(u) \geq \lambda\|\nabla u\|^2_{L^2(\Omega)} - \|g\|_{L^2(\Omega)}\,\|u\|_{L^2(\Omega)}.$$
By Poincare inequality, and Youngs inequality (C changes from line to line)
$$\mathcal{E}(u) \geq \lambda\|\nabla u\|^2_{L^2(\Omega)} - C\|g\|_{L^2(\Omega)}\,\|\nabla u\|_{L^2(\Omega)} \geq \lambda\|\nabla u\|^2_{L^2(\Omega)} - C\|g\|^2_{L^2(\Omega)} - \frac{\lambda}{2}\|\nabla u\|^2_{L^2(\Omega)}.$$

[50] $g \in H^{-1}(\Omega)$ suffices

That is,

$$\frac{\lambda}{2}\|\nabla u\|_{L^2(\Omega)}^2 \leq \mathcal{E}(u) + C\|g\|_{L^2}^2.$$

Again, by Poincaré inequality, this implies

$$\|u\|_{H^1(\Omega)}^2 \lesssim \mathcal{E}(u) + C\|g\|_{L^2}^2. \tag{22.95}$$

- Now let $u_k \in H^1_0(\Omega)$ be a sequence such that $\mathcal{E}(u_k) \xrightarrow{k\to\infty} \inf_{H^1_0(\Omega)} \mathcal{E}$. Since $0 \in H^1_0(\Omega)$ we may assume, w.l.o.g., that $\mathcal{E}(u_k) \leq \mathcal{E}(0) = 0$. By coercivity, (22.95),

$$\|u_k\|_{H^1(\Omega)}^2 \lesssim \mathcal{E}(u_k) + C\|g\|_{L^2}^2 \leq 0 + C\|g\|_{L^2}^2,$$

that is

$$\sup_{k\in\mathbb{N}} \|u_k\|_{H^1(\Omega)}^2 < \infty.$$

By weak compactness, Rellich, Theorem 13.25 and Theorem 13.35 we may assume w.l.o.g. (otherwise taking a subsequence) that there is $u \in H^1_0(\Omega)$ and

- $u_k \xrightarrow{k\to\infty} u$ strongly in $L^2(\Omega)$ and a.e.
- ∇u_k weakly converges to ∇u in $L^2(\Omega)$, that is for any $G \in L^2(\Omega,\mathbb{R}^n)$,

$$\int_\Omega \nabla u_k G \xrightarrow{k\to\infty} \int_\Omega \nabla u G.$$

Consequently,

$$\int_\Omega g u_k \xrightarrow{k\to\infty} \int_\Omega g u.$$

Moreover,

$$\begin{aligned}\int_\Omega a_{ij}\partial_i u_k \partial_j u_k &= \int_\Omega a_{ij}\partial_i u_k \partial_j u + \int_\Omega a_{ij}\partial_i u_k \partial_j (u_k - u)\\ &= \int_\Omega a_{ij}\partial_i u_k \partial_j u + \underbrace{\int_\Omega a_{ij}\partial_i u \partial_j (u_k - u)}_{\xrightarrow{k\to\infty} 0} + \underbrace{\int_\Omega a_{ij}\partial_i (u_k-u) \partial_j (u_k - u)}_{\geq 0}\end{aligned}$$

so

$$\liminf_{k\to\infty} \int_\Omega a_{ij}\partial_i u_k \partial_j u_k \geq \liminf_{k\to\infty} \int_\Omega a_{ij}\partial_i u_k \partial_j u = \int_\Omega a_{ij}\partial_i u \partial_j u \tag{22.96}$$

Together we have shown

$$\liminf_{k\to\infty} \mathcal{E}(u_k) \geq \mathcal{E}(u).$$

But since $u \in H^1_0(\Omega)$ we have that

$$\mathcal{E}(u) \geq \inf_{H^1_0(\Omega)} \mathcal{E},$$

and by construction

$$\lim_{k\to\infty} \mathcal{E}(u_k) = \inf_{H^1_0(\Omega)} \mathcal{E}$$

So (22.96) implies that

$$\mathcal{E}(u) = \inf_{H^1_0(\Omega)} \mathcal{E},$$

that is u minimizes $\mathcal{E}$ in $H^1_0(\Omega)$.

- we have found a minimizer of $\mathcal{E}$ and this minimizer satisfies the equation.

We still need to show uniqueness and the estimate (22.94) of the solution. Observe that uniqueness follows once we show that (22.94) holds for any solution $u \in H^1_0(\Omega)$ of (22.93). Assume this is true, then if we have two solutions u and v for the same g, then $w := u-v \in H^1_0(\Omega)$ satisfies

$$\partial_i(a_{ij}\partial_j w) = 0.$$

By (22.94) we have $\|w\|_{H^1} = 0$ that is $w \equiv 0$.

It remains to show (22.94) for an arbitrary solution $u \in H^1_0(\Omega)$ of (22.93). This follwos from multiplying (22.93) by u and integrating. By ellipticity we then find,

$$\|\nabla u\|^2_{L^2(\Omega)} \lesssim \int_\Omega a_{ij}\partial_i u\, \partial_j u = |\int_\Omega gu| \lesssim \|g\|_{L^2(\Omega)}\, \|u\|_{L^2(\Omega)}.$$

By Poincare on the left-hand side we find

$$\|u\|^2_{H^1(\Omega)} \leq |g\|_{L^2(\Omega)}\, \|u\|_{H^1(\Omega)}.$$

Dividing both sides by $\|u\|_{H^1(\Omega)}$ we have established (22.93). □

Now we need to take care of the lower order guys in (22.92), and for this we use a fixed point argument. Let f be fixed and for $u \in H^1_0(\Omega)$ set $Tu \in H^1_0(\Omega)$ be the solution of

$$\partial_i(a_{ij}\partial_j(T_f u)) = -\frac{1}{\lambda_0} f + (\frac{1}{\lambda_0} + \omega)u - b_j\partial_j u - cu. \tag{22.97}$$

By Lemma 22.25 is well-defined, linear, and bounded (observe if $u \in H^1$ then the right-hand side of the equation above belongs to $L^2(\Omega)$). But more is true: for any $f \in L^2(\Omega)$, T_f is compact, which will be a consequence of the following estimate

Lemma 22.26. *For $f \in L^2(\Omega)$ let $T_f : H^1_0(\Omega) \to H^1_0(\Omega)$ be defined as above. Then*

$$\|T_f(u)\|_{H^1(\Omega)} \leq C\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right)$$

and

$$\|T_f(u) - T_f(v)\|_{H^1(\Omega)} \leq C\, \|u-v\|_{L^2(\Omega)}$$

Proof. Testing (22.97) we have by ellipticity and Poincaré inequality

$$\|T_f(u)\|^2_{H^1(\Omega)} \lesssim C\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right)\, \|T_f(u)\|_{L^2(\Omega)} + \left|\int_\Omega b_j\partial_j u\, T_f(u)\right|$$

By another integration by parts, (recall: $b \in C^\infty$)

$$\left|\int_\Omega b_j\partial_j u\, T_f(u)\right| \leq C(b, \nabla b)\, \|u\|_{L^2(\Omega)}\|T_f(u)\|_{H^1(\Omega)}.$$

Consequently, we find

$$\|T_f(u)\|^2_{H^1(\Omega)} \lesssim C\left(\|u\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\right)\, \|T_f(u)\|_{H^1(\Omega)},$$

and dividing both sides by $\|T_f(u)\|_{H^1(\Omega)}$ we find the first estimate.

The second estimate follows by the same argument: observe that $T_f(u)-T_f(v) = T_0(u-v)$. Then, by the above argument,

$$\|T_0(u-v)\|^2_{H^1(\Omega)} \precsim C\,\|u-v\|_{L^2(\Omega)}$$

□

From Lemma 22.26 we readily conclude that T_f is actually compact for any $f \in L^2(\Omega)$

Corollary 22.27. *For $f \in L^2(\Omega)$, let T_f be defined as above. Then $T_f : H^1_0(\Omega) \to H^1_0(\Omega)$ is a continuous, compact operator. That is, for any bounded sequence $(u_k)_k \in H^1_0(\Omega)$,*

$$\sup_k \|u_k\|_{H^1(\Omega)} < \infty$$

there exists a subsequence u_{k_i} such that $(T_f u_{k_i})_i$ is convergent in $H^1_0(\Omega)$.

Proof. Clearly from Lemma 22.26 we obtain that T_f is Lipschitz continuous.

Now let u_k be a bounded H^1_0-sequence. Up to a subsequence we can assume that u_k weakly converge to some $u \in H^1_0(\Omega)$ and by Rellich, Theorem 13.35, the convergence is strong in $L^2(\Omega)$. By Lemma 22.26,

$$\|T_f u_k - T_f u\|_{H^1(\Omega)} \precsim \|u_k - u\|_{L^2(\Omega)} \xrightarrow{k\to 0} 0.$$

Thus, T_f is compact. □

Observe that this does not mean that T_f is contracting, so we cannot apply e.g. Banach Fixed Point argument. Instead we use the *Schauder fixed point theorem* in the form of *Schaefer's fixed point theorem* also known as the *Leray-Schauder theorem*, Corollary 18.10.

We can apply this theorem to the compact operator T essentially since we dissipativity:

Lemma 22.28. *There exist $\omega_0 > 0$ such that for any $\omega > \omega_0$ and any $\lambda_0 > 0$ we have the following:*

Let $\mu \in [0,1]$ and assume that $u \in H^1_0(\Omega)$ satisfies

$$u = \mu T u$$

then with a constant independent of μ,

$$\|u\|_{H^1(\Omega)} \leq C\|f\|_{L^2(\Omega)}$$

Lemma 22.29. *From $u = \mu T u$ we conclude in view of the definition of T, (22.97),*

$$-\partial_i(a_{ij}\partial_j u) = -\mu\left(-\frac{1}{\lambda_0}f + (\frac{1}{\lambda_0}+\omega)u - b_j\partial_j u - cu\right).$$

Similar to the proof of A being dissipativ, we test this equation with u, and obtain

$$\|\nabla u\|^2_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)} + \int_\Omega \left(c - \omega - \frac{1}{\lambda_0}\right)|u|^2 + \int_\Omega b_j\partial_j u\, u.$$

Observe that $\partial_j u u = \frac{1}{2}\partial_j |u|^2$ and integrating by parts we find

$$\|\nabla u\|^2_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)} - \left(\omega + \frac{1}{\lambda_0}\right)\int_\Omega |u|^2 + C(c,b)\,\|u\|^2_{L^2(\Omega)}$$

If ω is large enough, namely $\omega > C(c,b)$ this implies

$$\|\nabla u\|^2_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}.$$

By Poincaré inequality we obtain the claim.

No we can conclude:

A is m-dissipative. We argued above that A is m-dissipative if for any $f \in L^2(\Omega)$ there exists $u \in H^2(\Omega) \cap H^1_0(\Omega)$ such that $Tu = u$. Since T is compact by Corollary 22.27 and in view of Lemma 22.28, Leray-Schauder in form of Corollary 18.10 is applicable, and implies that there exists a fixed point $u \in H^1_0(\Omega)$ such that $Tu = u$.

Observe that this implies that u solves an equation of the form

$$\begin{cases} \partial_i(a_{ij}\partial_j u) = g \quad \text{in } \Omega \\ u = 0 \quad \text{in } \partial\Omega \end{cases}$$

for $g \in L^2(\Omega)$.

Now with PDE theory we can obtain interior regularity, indeed we find

$$\|u\|_{W^{2,2}_{loc}(\Omega)} \leq C(\Omega)\,\|g\|_{L^2(\Omega)}.$$

Close to the boundary one needs to do a reflection argument to obtain the same estimate up to the boundary.

Up to doing this, we have shown that there is $u \in D(A)$ and $(I - \lambda_0 A)u = f$;

That is, A is m-dissipative, and thus generates a contractive semigroup as claimed. □

Exercise 22.30. *Find a solution to the inhomogeneous equation*

$$\begin{cases} \partial_t u(x,t) - Lu(x,t) = f(x,t) \quad \textit{in } \Omega \times (0,\infty) \\ u(x,0) = u_0(x). \end{cases}$$

Here L is the usual elliptic operator with smooth coefficients.

Find an estimate of $\|u(t)\|_{L^2(\Omega)}$ in terms of $\|f\|_{L^2}, \|u_0\|_{L^2}$ and t.

hint: Use the Duhamel Ansatz from the heat equation in $\mathbb{R}^n$, i.e. set

$$u(t) = \int_0^t v_s(t)\,ds$$

where $v_s(t)$ is the solution associated to

$$\begin{cases} \partial_t v(x,t) - Lv(x,t) = 0 \quad \text{in } \Omega \times (s,\infty) \\ v(x,s) = f(x,s). \end{cases}$$

Observe that $v_s(t)$ can be written as semigroup!

Email address: armin@pitt.edu