ANALYSIS I & II

VERSION: April 18, 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, 2014a] (Fourier Analysis) and wikipedia. Sometimes we follow those sources verbatim.

Pictures that were not taken from above mentioned sources 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.

[Bourgain et al., 2001] Bourgain, J., Brezis, H., and Mironescu, P. (2001). Another look at Sobolev spaces. In *Optimal control and partial differential equations*, pages 439–455. IOS, Amsterdam.

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

[Brezis and Nguyen, 2011] Brezis, H. and Nguyen, H. (2011). The Jacobian determinant revisited. *Invent. Math.*, 185(1):17–54.

[Bucur and Valdinoci, 2016] Bucur, C. and Valdinoci, E. (2016). *Nonlocal diffusion and applications*, volume 20 of *Lecture Notes of the Unione Matematica Italiana*. Springer, [Cham]; Unione Matematica Italiana, Bologna.

[Bui and Candy, 2015] Bui, H.-Q. and Candy, T. (2015). A characterisation of the Besov-Lipschitz and Triebel-Lizorkin spaces using Poisson like kernels. *arXiv:1502.06836*.

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

[Di Nezza et al., 2012] Di Nezza, E., Palatucci, G., and Valdinoci, E. (2012). Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573.

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

[Grafakos, 2014b] Grafakos, L. (2014b). *Modern Fourier analysis*, volume 250 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.

[Lenzmann and Schikorra, 2019] Lenzmann, E. and Schikorra, A. (2019). Sharp commutator estimates via harmonic extensions. *Nonlinear Analysis (accepted)*.

[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.

[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.

[Mironescu and Sickel, 2015] Mironescu, P. and Sickel, W. (2015). A Sobolev non embedding. *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.*, 26(3):291–298.

[Runst and Sickel, 1996] Runst, T. and Sickel, W. (1996). *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *De Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin.

[Samko, 2002] Samko, S. G. (2002). *Hypersingular integrals and their applications*, volume 5 of *Analytical Methods and Special Functions*. Taylor & Francis, Ltd., London.

[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.

[Stein, 1993] Stein, E. M. (1993). *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[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.

[Tartar, 2007] Tartar, L. (2007). *An introduction to Sobolev spaces and interpolation spaces*, volume 3 of *Lecture Notes of the Unione Matematica Italiana*. Springer, Berlin; UMI, Bologna.

[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 $(x_i + O_i C_i)_{i=1}^N$ are pairwise disjoint and

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

  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.

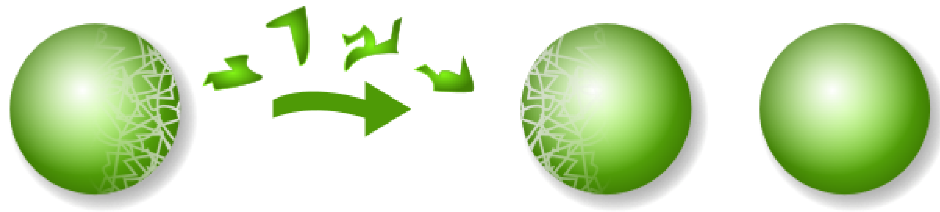


FIGURE 1.1. A ball can be decomposed into a finite number of disjoint sets and then reassembled into two balls identical to the original

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*

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

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

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

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

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

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*,

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

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 $\{q_2, \dots, \} = \mathbb{Q} \cap [0,1]$ 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 $0 \le s < t < \infty$.*

(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 $r_k \le \delta$ and $(B(x_k, r_k))_{k\in\mathbb{N}}$ cover A we have

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

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

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

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

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

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

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

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

So if $s > 2$ we see that $\mathcal{H}^s(C) \le \lim_{\delta\to 0} \delta^{s-2} = 0$. 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 other direction. 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.

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

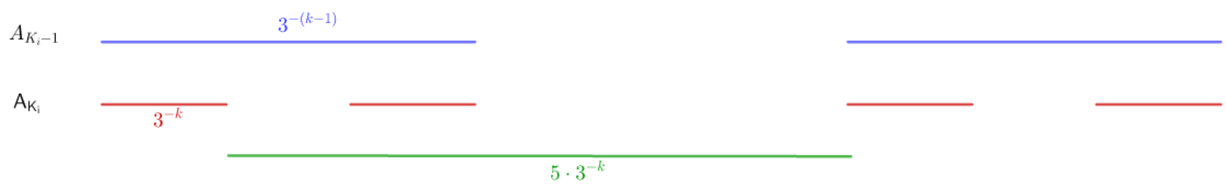


FIGURE 1.3. If a ball intersects three intervals of A_{K_i} its *diameter* is at least $5 \cdot 3^{-K_i}$

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

(2) *If $A_1 \subset A_2 \subset \ldots \subset A_k \subset \ldots$ then*

$$\mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \lim_{k \to \infty} \mu(A_k).$$

(3) *If $\mu(A_1) < \infty$ and $A_1 \supset A_2 \supset \ldots \supset A_k \supset \ldots$ then*

$$\mu\Big(\bigcap_{k=1}^{\infty} A_k\Big) = \lim_{k \to \infty} \mu(A_k).$$

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 $\mu(A_k) \le \mu(A_1) < \infty$)

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

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 $X\setminus N$ 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 $\mu(B\setminus N) \le \mu(B)$. So we have
$$\mu(B \cap N) + \mu(B\setminus N) \le \mu(N) + \mu(B) = \mu(B).$$
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

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

<u>Exercise: $\mathcal{A}$ is an Algebra</u>

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

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

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

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

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

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

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

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

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

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

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

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

Now we show $\lambda(A) \le \mu(A)$. 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

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

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

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

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

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

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

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

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

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

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

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 $\mathrm{dist}\,(A \cap G_k, A \backslash G) > 0$ 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

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

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

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

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 $\delta := \text{dist}\,(E, F) > 0$. 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

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

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. □

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

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

□

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 $\mathcal{S}$ is a π-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 $\mathcal{S}$ is a σ-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. □

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

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

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

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

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

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

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. □

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

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

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

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

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

$$\mathcal{H}^s_{\frac{1}{\ell}}\left(\bigcup_{k=1}^\infty B(r_{k;\ell})\right) \leq \mathcal{H}^s_{\frac{1}{\ell}}(E) + \frac{1}{\ell} \leq \mathcal{H}^s(E) + \frac{1}{\ell}. \tag{1.15}$$

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

$$\mathcal{H}^s(G) = \lim_{\ell\to\infty} \mathcal{H}^s_{\frac{2}{\ell}}(G) \leq \limsup_{\ell\to\infty} \mathcal{H}^s_{\frac{2}{\ell}}\left(\bigcup_{k=1}^\infty B(r_{k;\ell})\right) \overset{(1.15)}{\leq} \mathcal{H}^s(E) + 0$$

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) ⇒ (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) ⇒ (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) ⇒ (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.

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

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

(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 $\mu(H \backslash A) = 0$.

(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.

(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. □

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

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

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

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*

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

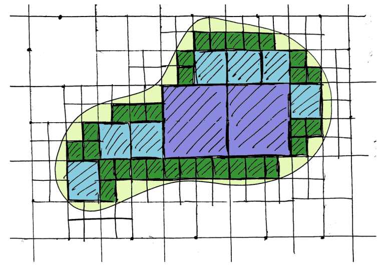


FIGURE 1.5. An open set can be represented by a countable union of dyadic cubes.

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

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

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 $\tilde{Q}$. □

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. $\square$

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 $E = B(0,1)$. $\square$

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.

$\Rightarrow$: 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 $f(B_{kl}(-s)) =$

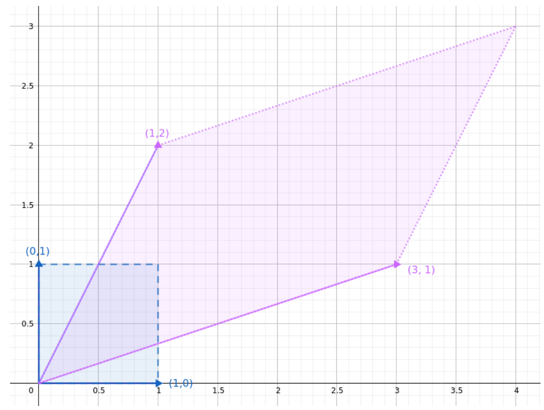


FIGURE 1.6. Let $a_{11} = 3$, $a_{12} = 1$, $a_{21} = 1$ and $a_{22} = 2$ then $[0,1]^2$ (blue) is transformed into $A[0,1]^2$ (purple). The purple area is $\det(A) = 5$.

$f(B_{kl}(s))$. 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. □

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

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

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. $\underline{(1) \Rightarrow (2)}$ 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.

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

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

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. □

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 $u_1, u_2, \ldots, u_n : X \to \mathbb{R}$ are measurable and $\Phi : \mathbb{R}^n \to Y$ is continuous, then the function

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

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 $f(x) \leq g(x)$ for all x and both f and g are simple functions, then
$$\int_\Omega f d\mu \leq \int_\Omega g d\mu.$$
- 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 $0 \leq s \leq f$.
- 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 $0 \leq s \leq g$ 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 $f_n(x) \le f(x)$ everywhere we have

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu \le \int_X f d\mu. \tag{3.1}$$

To conclude we need to show "$\ge$" in the above inequality.

Let $0 \le s \le f$ be a simple function. For a fixed constant $0 < c < 1$ we define

$$E_n := \{x \in X : \quad f_n(x) \ge c\, s(x)\}, \quad n \in \mathbb{N}.$$

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

$$\int_X f_n d\mu \ge \int_{E_n} f_n d\mu \ge c \int_{E_n} s d\mu = cs\llcorner\mu(E_n). \tag{3.2}$$

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

$$\alpha = \lim_{n\to\infty} \int_X f_n d\mu \ge c \int_X s d\mu.$$

Recall that this holds whenever s is a simple function with $s \le f$. Taking the supremum over such s we conclude

$$\alpha \ge c \int_X f d\mu.$$

This holds for any $c \in (0,1)$, taking the limit as $c \to 1$ we find

$$\alpha \ge \int_X f d\mu.$$

Combining this with (3.1) we conclude. □

Corollary 3.7 (Linearity of the integral)**.** *Let $f, g : X \to [0,\infty]$ μ-measurable and $\lambda, \sigma \ge 0$. 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*

$$\int_X \left(\liminf_{n\to\infty} f_n \right) d\mu \le \liminf_{n\to\infty} \int_X f_n d\mu.$$

Proof. Let $g_n := \inf\{f_n, f_{n+1}, \ldots\}$. Then $g_n \le f_n$ and hence

$$\int_X g_n d\mu \le \int_X f_n d\mu$$

Since $0 \le g_1 \le g_2 \le \ldots$, all the functions g_n are measurable, Theorem 2.13. Taking the $\liminf$ in the above inequality we have

$$\liminf_{n\to\infty} \int_X g_n d\mu \le \liminf_{n\to\infty} \int_X f_n d\mu.$$

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 $\lim_{n\to\infty} g_n = \liminf_{k\to\infty} f_k$. □

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*

$$\tilde{\nu}(E) := \int_E f \, d\mu \quad \forall E \text{ } \mu\text{-measurable}$$

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

$$\int_X g \, d\,(f \llcorner \mu) = \int_X gf \, d\mu. \tag{3.3}$$

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. □

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

[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 $f(x) = f(x)\chi_A(x) + f(x)\chi_{X\backslash A}(x)$

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.

□

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

$$\int_X f \ d\mu = \left(\int_X f^1 \ d\mu, \ldots, \int_X f^n \ d\mu \right).$$

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. □

[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 $\operatorname{supp} \varphi \subset U$, $0 \le \varphi \le 1$, $\varphi(x) = 1$ for $x \in K$ – for example if we set $d := \frac{1}{2}\operatorname{dist}(K, X \setminus U)$ we can choose $\varphi(x) := \frac{(d - \operatorname{dist}(x,K))_+}{d}$. We then have

$$\|\chi_E - \varphi\|_{L^p} = \left(\int_{X\setminus K} |\chi_E - \varphi|^p d\mu\right)^{\frac{1}{p}} \le \left(\int_{X\setminus K} |\chi_E|^p d\mu\right)^{\frac{1}{p}} + \left(\int_{X\setminus K} |\varphi|^p d\mu\right)^{\frac{1}{p}}$$
$$\le \mu(E \setminus K)^{1/p} + \mu(U \setminus K)^{1/p} < \varepsilon.$$

□

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 $1 \le p < \infty$, 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 $\operatorname{supp} f$ *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 $f_t := \chi_{[0,t]}$, $0 < t \le 1$. 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

$$\int_{[0,1]} f(x)d\mathcal{L}^1(x) \le \sup_{[0,1]} f \ \mathcal{L}^1([0,1]) < \infty.$$

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

$$\sum_{i=1}^N (x_i - x_{i-1})M_i - \varepsilon \le \mathrm{Ri} \int_{[0,1]} f(x)dx \le \sum_{i=1}^N (x_i - x_{i-1})m_i + \varepsilon$$

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

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

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. □

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 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. □

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. $\Rightarrow$ 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.

$\Leftarrow$ 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 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. □

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 $x \notin \bigcup_{i=k}^{\infty} \{x : |f_{n_i}(x) - f(x)| \geq \frac{1}{i}\}$. 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

$$\begin{aligned}\mu(\{x : f(x) > \lambda\}) =& \int_{\{x:f(x)>\lambda\}} d\mu\\ =&\lambda^{-p} \int_{\{x:f(x)>\lambda\}} \lambda^p d\mu\\ \le&\lambda^{-p} \int_{\{x:f(x)>\lambda\}} |f|^p d\mu.\end{aligned}$$

□

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 $\Lambda \ge 0$ such that for all $\lambda > 0$

$$\mu(\{x : f(x) > \lambda\}) \le \frac{1}{\lambda^p}\Lambda^p.$$

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 $\|f\|_{L^{(p,\infty)}} \le \|f\|_{L^p}$. 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

$$\mu\{|f_k - f| > \varepsilon\} \le \varepsilon^{-p}\|f_k - f\|^p_{L^p} \xrightarrow{k\to\infty} 0.$$

□

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

$$\left|\|f_k\|_{L^p(X)} - \|f\|_{L^p(X)}\right| \le \|f_k - f\|_{L^p(X)} \xrightarrow{k\to\infty} 0.$$

(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 $1 \le p < \infty$)

$$2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p \ge 0.$$

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,

$$\begin{aligned}
&2^p \int_\Omega |f|^p d\mu - \limsup_{k\to\infty} \int_\Omega |f_k - f|^p d\mu \\
=&\liminf_{k\to\infty} \int_\Omega \left(2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p\right) d\mu \\
\ge&\int_\Omega \liminf_{k\to\infty} \left(2^{p-1}|f_k|^p + 2^{p-1}|f|^p - |f_k - f|^p\right) d\mu \\
=&2^p \int_\Omega |f|^p d\mu
\end{aligned}$$

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) $\Rightarrow$ (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) $\Rightarrow$ (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). $\square$

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

$$(\mu \times \nu)(S) := \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i)\nu(B_i) : \quad S \subset \bigcup_{i=1}^{\infty} A_i \times B_i; \quad A_i \subset X \text{ } \mu\text{-measurable, } B_i \subset Y \text{ } \nu\text{-measurable} \right\}$$

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. □

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*

$$\int_{(0,R)} |f(x,r)| dx \le \frac{1}{R} \int_{(0,R)^2} |f(x,y)| d(x,y).$$

More precisely, consider for $\lambda > 1$ the set $Y_\lambda \subset \mathbb{R}$

$$Y_\lambda := \left\{ r \in (0,R) : \int_{(0,R)} |f(x,r)| dx \le \frac{\lambda}{R} \int_{(0,R)^2} |f(x,y)| d(x,y) \right\}.$$

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

$$\mathcal{L}^1(X_\lambda) \frac{\lambda}{R} \int_{(0,R)^2} |f(x,y)| d(x,y) \leq \int_{X_\lambda} \int_{(0,R)} |f(x,r)| dx dr \leq \int_{(0,R)^2} |f(x,y)| dx dy$$

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

$$T(\lambda f + \mu g) = \lambda T f + \mu T g.$$

(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

$$\lambda \mapsto \lambda^{p-1} \mathcal{L}^n \left(\{x \in \Omega : |f(x)| > \lambda\} \right) \quad \text{in } L^1((0, \infty)).$$

Moreover, in either case

$$\int_\Omega |f(x)|^p dx = p \int_{\lambda=0}^\infty \lambda^{p-1} \mathcal{L}^n \left(\{x \in \Omega : |f(x)| > \lambda\} \right) d\lambda.$$

[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 $|Tf| \le |Tf_1| + |Tf_2|$,

$$\{|Tf| > \lambda\} \subset \{|Tf_1| > \frac{\lambda}{2}\} \cup \{|Tf_2| > \frac{\lambda}{2}\}$$

so

$$\mathcal{L}^n\left(\{|Tf| > \lambda\}\right) \le \mathcal{L}^n\left(\{|Tf_1| > \frac{\lambda}{2}\}\right) + \mathcal{L}^n\left(\{|Tf_2| > \frac{\lambda}{2}\}\right).$$

By Chebyshev inequality, Lemma 3.48, for $i = 1, 2$ and boundedness of T,

$$\begin{aligned} \mathcal{L}^n\left(\{|Tf_i| > \frac{\lambda}{2}\}\right) &\le \left(\frac{\lambda}{2}\right)^{-p_i} \|Tf_i\|_{L^{p_i}(\mathbb{R}^n)}^{p_i} \\ &\le \left(\frac{\lambda}{2}\right)^{-p_i} \|f_i\|_{L^{p_i}(\mathbb{R}^n)}^{p_i} \end{aligned}$$

By the definition of f_i we have

$$\left(\frac{\lambda}{2}\right)^{-p_1} \|f_1\|_{L^{p_1}(\mathbb{R}^n)}^{p_1} = \left(\frac{\lambda}{2}\right)^{-p_1} \int_{\{|f(\cdot)|>\lambda\}} |f(x)|^{p_1} dx$$

$$\left(\frac{\lambda}{2}\right)^{-p_2} \|f_2\|_{L^{p_2}(\mathbb{R}^n)}^{p_2} = \left(\frac{\lambda}{2}\right)^{-p_2} \int_{\{|f(\cdot)| \leq \lambda\}} |f(x)|^{p_2} dx$$

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

$$\begin{aligned}\|Tf\|_{L^p(\mathbb{R}^n)}^p \leq & p \int_{\lambda=0}^{\infty} \lambda^{p-1} \mathcal{L}^n\left(\{x \in \Omega : |Tf(x)| > \lambda\}\right) d\lambda \\ \leq & C \int_{\lambda=0}^{\infty} \lambda^{p-1} \left(\lambda^{-p_1} \int_{\{|f(\cdot)| > \lambda\}} |f(x)|^{p_1} dx + \lambda^{-p_2} \int_{\{|f(\cdot)| \leq \lambda\}} |f(x)|^{p_2} dx\right) d\lambda\end{aligned}$$

Again by Fubini

$$\begin{aligned}\|Tf\|_{L^p(\mathbb{R}^n)}^p \leq & C \int_{\mathbb{R}^n} |f(x)|^{p_1} \left(\int_{\lambda=0}^{|f(x)|} \lambda^{p-1} \lambda^{-p_1} d\lambda\right) dx \\ & + C \int_{\mathbb{R}^n} |f(x)|^{p_2} \left(\int_{\lambda=|f(x)|}^{\infty} \lambda^{p-1} \lambda^{-p_2} d\lambda\right) dx\end{aligned}$$

Since $p_1 < p < p_2$ the integrals in λ converge and we have

$$\begin{aligned}\|Tf\|_{L^p(\mathbb{R}^n)}^p \leq & C \int_{\mathbb{R}^n} |f(x)|^{p_1} \left(|f(x)|^p |f(x)|^{-p_1}\right) dx \\ & + C \int_{\mathbb{R}^n} |f(x)|^{p_2} \left(|f(x)|^p |f(x)|^{-p_2} d\lambda\right) dx \\ = & \tilde{C} \int_{\mathbb{R}^n} |f(x)|^p dx \\ = & \tilde{C} \|f\|_{L^p(\mathbb{R}^n)}^p.\end{aligned}$$

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, 2014a, §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*
$$T(\lambda f + \mu g) = \lambda T f + \mu T g.$$

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

$$\begin{aligned}\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy \leq &\left(\int_{\mathbb{R}^n} |f(x-y)|^p dy\right)^{\frac{1-\frac{p}{r}}{p}} \left(\int_{\mathbb{R}^n} |g(y)|^q dy\right)^{\frac{1-\frac{q}{r}}{q}} \left(\int_{\mathbb{R}^n} |f(x-y)|^p |g(y)|^q dy\right)^{\frac{1}{r}}\\ =&\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{1-\frac{q}{r}} \left(|f|^p * |g|^q\right)^{\frac{1}{r}}\end{aligned}$$

Now from the L^1-case,

$$\begin{aligned}\| \left(|f|^p * |g|^q\right)^{\frac{1}{r}} \|_{L^r(\mathbb{R}^n)} =&\| |f|^p * |g|^q\|_{L^1(\mathbb{R}^n)}^{\frac{1}{r}} \leq \left(\| |f|^p\|_{L^1(\mathbb{R}^n)} \| |g|^q\|_{L^1(\mathbb{R}^n)}\right)^{\frac{1}{r}}\\ =&\|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{\frac{q}{r}}\end{aligned}$$

Plugging all this together we find

$$\begin{aligned}\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x-y)g(y)|\, dy\right)^r dx \leq &\left(\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{1-\frac{q}{r}} \| \left(|f|^p * |g|^q\right)^{\frac{1}{r}} \|_{L^r(\mathbb{R}^n)}\right)^r\\ \leq &\left(\|f\|_{L^p(\mathbb{R}^n)}^{1-\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{r-\frac{q}{r}} \|f\|_{L^p(\mathbb{R}^n)}^{\frac{p}{r}} \|g\|_{L^q(\mathbb{R}^n)}^{\frac{q}{r}}\right)^r\\ = &\left(\|f\|_{L^p(\mathbb{R}^n)} \|g\|_{L^q(\mathbb{R}^n)}\right)^r\end{aligned}$$

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. □

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.

□

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,

$$\sup_{x\in K} |f(x) - f_\varepsilon(x)| = \sup_{x\in K} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\eta_\varepsilon(y)dy \right|$$
$$\leq \sup_{x\in K} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|\, \eta_\varepsilon(y)dy \to 0 \quad \text{as } \varepsilon \to 0.$$

(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 $f \in L^p(\mathbb{R}^n)$, $1 \leq p \leq \infty$ then $f_\varepsilon \in L^p(\mathbb{R}^n)$ and $\|f_\varepsilon\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)}$.*

Proof. Follows from Theorem 4.18:

$$\|f_\varepsilon\|_{L^p(\mathbb{R}^n)} = \|f * \eta_\varepsilon\|_{L^p(\mathbb{R}^n)} \leq \|f\|_{L^p(\mathbb{R}^n)} \underbrace{\|\eta_\varepsilon\|_{L^1(\mathbb{R}^n)}}_{=1}.$$

□

Theorem 4.31. *If $f \in L^p(\mathbb{R}^n)$ and $1 \leq p < \infty$ 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 $f \in L^p(\mathbb{R}^n)$, $1 \leq p < \infty$ 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

$$\|\tau_y f - f\|_{L^p} \leq \|\tau_y f - \tau_y g\|_{L^p} + \|\tau_y g - g\|_{L^p} + \|f-g\|_{L^p} = 2\|f-g\|_{L^p} + \|\tau_y g - g\|_{L^p}$$
$$< 2\varepsilon/3 + \|\tau_y g - g\|_{L^p}$$

Since $\tau_y g \to g$ converges uniformly, and $K := \bigcup_{|y|\le 1} \operatorname{supp} \tau_y g$ is bounded, there is $\delta > 0$ such that

$$\|\tau_y g - g\|_{L^p(\mathbb{R}^n)} = \|\tau_y g - g\|_{L^p(K)} \overset{\text{Hölder}}{\le} C(K)\, \|\tau_y g - g\|_{L^\infty(\mathbb{R}^n)} < \varepsilon/3 \quad \text{for } |y| < \delta.$$

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

$$\begin{aligned}
\|f - f_\varepsilon\|_{L^p} &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} (f(x) - f(x-y))\eta_\varepsilon(y) dy \right|^p dx \\
&\le \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(x) - f(x-y)| \eta_\varepsilon(y)^{1/p} \eta_\varepsilon(y)^{1/q} dy \right)^p dx \\
&\le \int_{\mathbb{R}^n} \left(|f(x) - f(x-y)|^p \eta_\varepsilon(y) dy \right) \underbrace{\left(\int_{\mathbb{R}^n} \eta_\varepsilon(y) dy \right)^{\frac{p}{q}}}_{=1} dx \\
&= \int_{\mathbb{R}^n} \int_{B(0,\varepsilon)} |f(x) - f(x-y)|^p \eta_\varepsilon(y) dy dx \\
&= \int_{B(0,\varepsilon)} \left(\int_{\mathbb{R}^n} |f(x) - f(x-y)|^p dx \right) \eta_\varepsilon(y)\, dy \xrightarrow{\varepsilon \to 0} 0.
\end{aligned}$$

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 $1 \le p < \infty$. 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. □

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

$$\|f_k\|_{L^p(\mathbb{R}^n)} \le \|f - f_k\|_{L^p(\mathbb{R}^n)} + \|f\|_{L^p(\mathbb{R}^n)}.$$

We also get one part of the inequality

$$\inf_{(\tilde{f}_k)_k} \liminf_{k\to\infty} \|\tilde{f}_k\|_{L^p(\mathbb{R}^n)} \le \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)} \le \liminf_{k\to\infty} \|f - f_k\|_{L^p(\mathbb{R}^n)} + \liminf_{k\to\infty} \|f\|_{L^p(\mathbb{R}^n)} = 0 + \|f\|_{L^p(\mathbb{R}^n)}.$$

$\underline{\Leftarrow}$ 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

$$\int_{\mathbb{R}^n} |f|^p = \int_{\mathbb{R}^n} \lim_{k\to\infty} |f_k|^p \le \liminf_{k\to\infty} \int_{\mathbb{R}^n} |f_k|^p.$$

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

$$\|f\|_{L^p(\mathbb{R}^n)} \le \inf_{(f_k)_k} \liminf_{k\to\infty} \|f_k\|_{L^p(\mathbb{R}^n)}$$

□

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 id $: \mathcal{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ *given by* $\operatorname{id} f := f$ *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 id $: \tilde{L}^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$, given by $\operatorname{id} f := \lim_{k\to\infty} \operatorname{id} f_k$ 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 $\chi_{[0,1]} \not\in W^{1,p}(\mathbb{R})$ 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 $\mathcal{G}_1, \ldots, \mathcal{G}_{N_n}$ 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]. □

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
$$\nu(A \cap V) \leq \mu(A \cap V) \quad \text{for all bounded Borel sets } V \subset \mathbb{R}^n$$
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 $\varepsilon \to 0$. □

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

$$\nu(A\cap V) = \nu(A\cap V\cap U) \le \sum_{B\in\mathcal{G}} \nu(B)+0 \le \sum_{B\in\mathcal{G}}(\alpha+\varepsilon)\mu(B)+0 = (\alpha+\varepsilon)\mu(\bigcup_{B\in\mathcal{G}} B) \le (\alpha+\varepsilon)\mu(U).$$

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 $\alpha > 0$

$$A \subset \{x \in \mathbb{R}^n : \overline{D}_\mu\nu(x) \ge \alpha\},$$

and we can apply Lemma 5.6(2) to obtain

$$\mu(A) \le \frac{1}{\alpha}\nu(A).$$

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

$$b\mu(R(a,b)) \le \nu(R(a,b)) \le a\mu(R(a,b)).$$

[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

$$f_E \equiv \fint_E d\mu := \mu(E)^{-1}\int_E f d\mu$$

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. $x \in \mathbb{R}^n$*

$$\lim_{r\to 0}\fint_{B(x,r)} f\, d\mu = f(x)$$

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

$$\fint_{B(x,r)} f d\mu = \frac{\nu_+(B(x,r))}{\mu(B(x,r))} - \frac{\nu_-(B(x,r))}{\mu(B(x,r))}$$
$$\xrightarrow{r\to 0^+} D_\mu \nu^+(x) - D_\mu \nu^-(x) = f^+(x) - f^-(x) = f(x),$$

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 $\mu(N) = 0$. □

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

$$S \subset B(x, r_S), \quad \mathcal{L}^n(B(x, r_S)) \le C\mathcal{L}^n(S),$$

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

$$\begin{aligned}
\left|\fint_S f d\mathcal{L}^n - f(x)\right| \le & \fint_S |f(y) - f(x)| d\mathcal{L}^n(y) \\
\le & \mathcal{L}^n(S)^{-1} \int_{B(x,r_s)} |f(y) - f(x)| d\mathcal{L}^n(y) \\
= & \frac{\mathcal{L}^n(B(x, r_S))}{\mathcal{L}^n(S)} \fint_{B(x,r_S)} |f(y) - f(x)| d\mathcal{L}^n(y) \\
\le & C \fint_{B(x,r_S)} |f(y) - f(x)| d\mathcal{L}^n(y) \xrightarrow{\mathcal{L}^n(S)\to 0} 0.
\end{aligned}$$

□

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. □

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
$$\mu \llcorner f(A) := \int_A f d\mu \quad A\ \mu\text{-measurable}$$
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

$$A_1 \subset A_2 \subset A_3 \subset \ldots$$

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*

$$L(X,Y) := \{T : X \to Y : \quad T \text{ linear and continuous}\}$$

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

$$(\lambda T + \mu S)\, x := \lambda T x + \mu S x.$$

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,
$$\|f_k\|^p_{L^p(\mathbb{R}^n)} = \|\sum_{i=1}^k \chi_{A_i\cap K}\|_{L^1(\mathbb{R}^n)} \xrightarrow{k\to\infty} \|f\|_{L^1(\mathbb{R}^n)} = \|f\|^p_{L^p(\mathbb{R}^n)}.$$
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

$$L(f) = \int_{\mathbb{R}^n} gf \, d\mathcal{L}^n(\mathbb{R}^n) \quad \forall f \in L^1(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n): \quad \text{supp} f \text{ compact}.$$

(simply take K to contain $\text{supp} f$).

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

$$\int_{\Phi(X)} f dx = \int_X f \circ \Phi |\det(D\Phi(x))|dx.$$

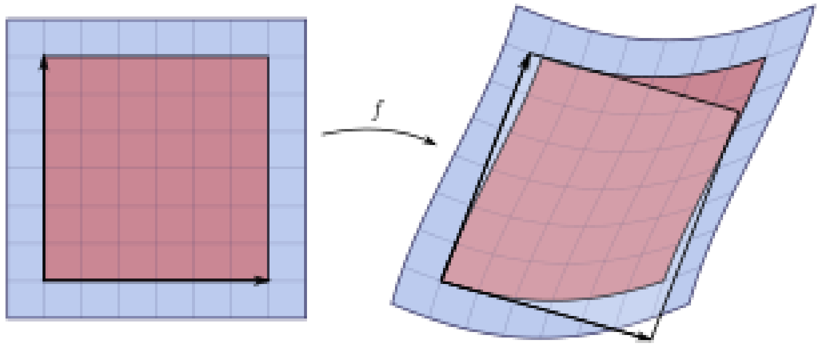


FIGURE 6.1. A nonlinear map $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ sends a small square (left, in red) to a distorted parallelogram (right, in red). The Jacobian at a point gives the best linear approximation of the distorted parallelogram near that point (right, in translucent white), and the Jacobian 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 $\Phi \in C^1$

$$|\Phi(z) - \Phi(x) - D\Phi(x)(z-x)| \leq o(r).$$

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 $\mathcal{L}^n(B(x,R)) = c_n R^n$, 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

$$|\mathrm{Jac}(\Phi)(z)| := \sqrt{\Big|\det_{\mathbb{R}^{n\times n}} (\underbrace{D\Phi^t(z)}_{\mathbb{R}^{n\times N}} \underbrace{D\Phi(z)}_{\mathbb{R}^{N\times n}})\Big|} \equiv \sqrt{\Big|\det_{\mathbb{R}^{n\times n}} \Big((\langle \partial_\alpha \Phi(z), \partial_\beta \Phi(z) \rangle_{\mathbb{R}^N})_{\alpha,\beta=1,\dots,n} \Big)\Big|}$$

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. □

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 $\mathbb{R}^n$,

$$\mathcal{H}^n(LA) = \mathcal{L}^n(SA) = |\det(S)| \mathcal{L}^n(A) = \sqrt{\det(L^t L)} \mathcal{L}^n(A).$$

$\square$

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, 2014a]. 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

$$\text{Topological spaces} \supset \text{metric spaces} \supset \text{normed vector spaces} \supset \text{Pre-Hilbert-spaces}$$

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 $Df = (\partial_1 f, \partial_2 f, \dots, \partial_n f) \in L^p(\Omega, \mathbb{R}^n)$.

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, $X = \operatorname{span}\{b_1, \dots, b_n\}$ for some $n \in \mathbb{N}$ and some $b_1, \dots, b_n \in X$, and $(b_i)_{i=1}^n$ *linear independent*, i.e. whenever $(\lambda_i)_{i=1}^n \subset \mathbb{R}$,

$$\sum_{i=1}^n \lambda_i b_i = 0 \quad \Rightarrow \quad \lambda_i = 0 \quad \forall i = 1, \dots, n.$$

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 seqeunce 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

$$\Lambda^{-1}\|x\|_2 \leq \|x\| \leq \Lambda\|x\|_2 \quad \forall x \in X.$$

Hint: To see $\Lambda^{-1}\|x\|_2 \leq \|x\|$ 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 $\|c_i\| \leq 1$ and*

$$\|c_i - c_j\| \geq \frac{1}{100} \quad \forall i \neq j.$$

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 (find the statement and proof online) to apply Exercise 7.16.

Exercise 7.18. *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

$$\|S \circ T\|_{\mathcal{L}(X,Z)} \le \|S\|_{\mathcal{L}(Y,Z)} \|T\|_{\mathcal{L}(X,Y)}.$$

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 $1 \leq p < \infty$ 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 $L^p(\Omega)$.

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).

[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

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 $(L^p(\Omega))^{n+1} = L^p(\Omega) \times \dots L^p(\Omega)$.

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

$$Tf := (f, \partial_1 f, \partial_2 f, \dots, \partial_n f).$$

Clearly $T : (W^{1,p}(\Omega), \|\cdot\|_{W1,p(\Omega)}) \to (L^p(\Omega))^{n+1}$ is injective. Moreover

$$\|Tf\|_{L^p(\Omega)\times\dots L^p(\Omega)} = \|f\|_{W^{1,p}(\Omega)}.$$

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 $\alpha \in \{1, \dots, n\}$. 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 $\tilde{T}^e : \mathrm{span}(v, W) \to \mathbb{R}$ which satisfies all the properties of an element in A. But since T^e is a maximal element we have $W \supset \mathrm{span}(v, W)$ – a contradiction since $v \not\in W$. We conclude that $W = X$ and T^e is the extension we wanted. □

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

$$Y := \text{span}\{x_0\}$$

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).

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

$$\sup_{x^* \in X^*, \|x^*\|_{X^*}=1} |x^*[x]| \leq \|x\|_X.$$

Now take $\bar{x}^*$ from Corollary 10.11. Then

$$\sup_{x^* \in X^*, \|x^*\|_{X^*}=1} |x^*[x]| \geq |\bar{x}^*[x]| = \|x\|_X.$$

[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.

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*

$$\|x\|_X = \sup_{u^*\in U, \|u^*\|_{X^*}\leq 1} |u^*[x]|$$

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*

$$\|f\|_{L^p(\mathbb{R}^n)} = \sup_{g\in C_c^\infty(\mathbb{R}^n), \|g\|_{L^{p'}}\leq 1} \int fg$$

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 $C = \left(\inf_{\tilde{u} \in U} \|\tilde{u} - x_0\|\right)^{-1}$. Thus $T \in V^*$ and we can use the Hahn-Banach extension to conclude. □

Recall the definition of separable spaces

Definition 10.19. A normed space X is *separable* iff there exists a countable set $U = \{x_1, x_2, \dots\}$ 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 $x_0^*(x_0) = 1$. □

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,

$$m_C(x+y) \le t+s.$$

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

$$m_C(x_0) \ge 1, \quad \text{and} \quad m_C(x) < 1 \quad \forall x \in C.$$

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 $m_C(x) \le t < 1$.

To see $m_C(x_0) \ge 1$ 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 $m_C(x_0) \ge 1$ (observe it could indeed be $= 1$ if x_0 lies on the boundary of C!).

Now set $Y := \text{span}\{x_0\}$ and define $y^* \in Y^*$ as

$$y^*[\lambda x_0] := \lambda m_C(x_0), \quad \lambda \in \mathbb{R}.$$

Since $m_C(x) \ge 0$ we find that

$$y^*[\lambda x_0] \begin{cases} = m_C(\lambda x_0) \quad \text{if } \lambda > 0 \\ \le 0 \le m_C(\lambda x_0) \text{if } \lambda \le 0 \end{cases}$$

Thus $y^*[y] \le m_C(y)$ for all $y \in Y$. By the Hahn-Banach theorem we find $x^* : X \to \mathbb{R}$ with

$$x^* : X \to \mathbb{R} \quad \text{linear, and}$$
$$x^*[x] \le m_C(x) \quad \forall x \in X.$$

In particular, in view of (10.5) we have that $x^* \in X^*$.

Thus we have

$$x^*[x] \le m_C(x) < 1 \quad \forall x \in C$$

and

$$x^*[x_0] = y^*[x_0] = m_C(x_0) \ge 1.$$

□

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*

$$x^*[u] < \lambda \le x^*[v] \quad \forall u \in U\,, v \in V.$$

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. □

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 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. □

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.

(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 $\|x_n - x\|_X \xrightarrow{n\to\infty} 0$.

[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

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 $x^\ast \in X^\ast$

$$x^\ast[x] \xleftarrow{k\to\infty} x^\ast[x_k] \xrightarrow{k\to\infty} x^\ast[y]$$

That is

$$x^\ast[x] = x^\ast[y],$$

so

$$x^\ast[x-y] = 0.$$

This holds for any $x^\ast$, 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 $X^\ast$ and weak*-convergence in $X^\ast$ 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*

$$\|x\|_X \leq \liminf_{k\to\infty} \|x_k\|_X$$

(2) *the dual norm is lower semicontinuous under weak* convergence. That is, assume $x_k^\ast$ weak*-converges to $x^\ast$. Then*

$$\|x^\ast\|_{X^\ast} \leq \liminf_{k\to\infty} \|x_k^\ast\|_{X^\ast}$$

Proof. (1) Assume x_k weakly converge to x. By Hahn-Banach, Corollary 10.14, there exists $x^\ast \in X^\ast$, $\|x^\ast\|_{X^\ast} = 1$, such that

$$\|x\|_X = x^\ast[x] = \lim_{k\to\infty} x^\ast[x_k] \leq \underbrace{\|x^\ast\|_{X^\ast}}_{\leq 1} \liminf_{k\to\infty} \|x_k\|_X.$$

(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

$$f_k[\varphi] \xrightarrow{k\to\infty} f[\varphi] \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

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

$$\|f_k\|_{L^q(\mathbb{R}^n)} = k^{\frac{n}{p}-\frac{n}{q}}\|f\|_{L^q(\mathbb{R}^n)} \le k^{\frac{n}{p}-\frac{n}{q}} C(\operatorname{supp} f)\|f\|_{L^p(\mathbb{R}^n)} \xrightarrow{k\to\infty} 0.$$

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

$$\int_{\mathbb{R}^n} g(f_k - f) + \int_{\mathbb{R}^n} G \cdot D(f_k - f) \xrightarrow{k\to\infty} 0 \quad \forall g, G \in L^{p'}. \tag{12.2}$$

$\Rightarrow$ 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).

$\Leftarrow$ 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 $f \in W^{1,p}(\mathbb{R}^n)$. □

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

$$\|T_1 x_1\| \geq \frac{2}{3} 3^{-1}\, \|T_1\|.$$

Repeating this inductively we find $x_n \in X$ such that $\|x_n - x_{n-1}\|_X \leq 3^{-n}$ and

$$\|T_n x_n\| \geq \frac{2}{3} 3^{-n} \|T_n\|.$$

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

$$\|T_n x\| \geq \|T_n x_n\| - \|T_n(x-x_n)\| \geq \frac{2}{3}3^{-n}\|T_n\| - \frac{1}{2}3^{-n}\|T_n\| = \frac{1}{6}3^{-n}\|T_n\| \geq \frac{1}{6}\left(\frac{4}{3}\right)^n.$$

That is,

$$\sup_{n\in\mathbb{N}} \|T_n x\|_X = +\infty,$$

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

$$(J_X x_n)[x^*] = x^*[x_n] \xrightarrow{n\to\infty} x^*[x] \quad \text{in } \mathbb{R}$$

Since convergent sequences in $\mathbb{R}$ are bounded, we find that for any $x^* \in X^*$

$$\sup_{n\in\mathbb{N}} |(J_X x_n)[x^*]| < \infty.$$

Since X^* is a Banach space we can apply Banach-Steinhaus Theorem 12.18 and find that actually

$$\sup_{n\in\mathbb{N}} \|J_X x_n\|_{X^{**}} < \infty,$$

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 $x_n \xrightarrow{*} x$. Show that there exists a sequence $(y_n)_{n\in\mathbb{N}}$ of convex combinations*

$$y_n = \sum_{k=1}^{N_n} \lambda_{n,k} x_k, \quad \text{with } \sum_{k=1}^{N_n} \lambda_{n,k} = 1,\ \lambda_{n,k} \in [0,1],\ N_n \in \mathbb{N}$$

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*

$$\Delta u - \lambda u = f \quad in\ \mathbb{R}^n$$

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

$$\frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu \geq (\frac{\lambda}{2} - \varepsilon^2)\|u\|_{L^2}^2 - \frac{1}{\varepsilon^2}\|f\|_{L^2(\mathbb{R}^n)}^2,$$

so if $\varepsilon^2 < \frac{\lambda}{2}$ we have

$$E(u) \geq \frac{\lambda}{2}\int_{\mathbb{R}^n} |u|^2 + \int_{\mathbb{R}^n} fu \geq -\frac{1}{\varepsilon^2}\|f\|_{L^2(\mathbb{R}^n)}^2 > -\infty.$$

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) $\limsup_{k\to\infty} \mu_k(K) \le \mu(K)$ for all compact sets $K \subset \mathbb{R}^n$ and $\mu(U) \le \liminf_{k\to\infty} \mu_k(U)$ 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. (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

$$\mu(K) = \inf_{K\subset U,\ U \text{ open}} \mu(U) \geq \limsup_{k\to\infty} \mu_k(K).$$

(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 $f \in C^0_c(\mathbb{R}^n)$. 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,

$$0 = t_0 < t_1 < \ldots < t_N \le 2$$

such that $|t_i - t_{i-1}| < \varepsilon$ and $t_N \ge \|f\|_{L^\infty}$ – and we may also assume $\mu(f^{-1}(t_i)) = 0$, whenever $i \ge 1$. This is possible by Exercise 1.74. Since $\|f\|_{L^\infty} \le 1$ we then have

$$\sum_{i=1}^N t_{i-1}\mu(f^{-1}[t_{i-1},t_i] \cap \overline{B}) \le \int_{\mathbb{R}^n} f d\mu \le \sum_{i=1}^N t_i \mu(f^{-1}[t_{i-1},t_i] \cap \overline{B})$$

and

$$\sum_{i=1}^N t_{i-1}\mu_k(f^{-1}[t_{i-1},t_i] \cap \overline{B}) \le \int_{\mathbb{R}^n} f d\mu_k \le \sum_{i=1}^N t_i \mu_k(f^{-1}[t_{i-1},t_i] \cap \overline{B})$$

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

$$\mu(\partial B_i) \le \mu(\partial \overline{B}) + \mu(\{t_{i-1}\}) + \mu(\{t_i\}).$$

Subtracting both inequalities we find

$$\int_{\mathbb{R}^n} f d\mu - \int_{\mathbb{R}^n} f d\mu_k \le \sum_{i=1}^N t_i \mu(f^{-1}[t_{i-1},t_i] \cap \overline{B}) - \sum_{i=1}^N t_{i-1}\mu_k(f^{-1}[t_{i-1},t_i] \cap \overline{B})$$

and

$$\int_{\mathbb{R}^n} f d\mu_k - \int_{\mathbb{R}^n} f d\mu \le \sum_{i=1}^N t_i \mu_k(f^{-1}[t_{i-1},t_i] \cap \overline{B}) - \sum_{i=1}^N t_{i-1}\mu(f^{-1}[t_{i-1},t_i] \cap \overline{B})$$

Taking the limit, using condition (2), we find that

$$\begin{aligned}
\limsup_{k\to\infty} \left| \int_{\mathbb{R}^n} f d\mu - \int_{\mathbb{R}^n} f d\mu_k \right| \le & \sum_{i=1}^N (t_i - t_{i-1})\, \mu(f^{-1}[t_{i-1},t_i] \cap \overline{B}) \\
\le & \varepsilon \sum_{i=1}^N \mu(f^{-1}[t_{i-1},t_i] \cap \overline{B}) \\
\le & 2\varepsilon \mu(\overline{B})
\end{aligned}$$

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. □

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 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 $1 \le p \le \infty$, $k \in \mathbb{N}$ 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!)

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\gamma|\leq k} \|\partial^\gamma u\|_{L^p(\mathbb{R}^n)}^p\right)^{\frac{1}{p}}.$$

(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 $W_{loc}^{k,p}(\Omega)$ is similarly defined as $L_{loc}^p(\Omega)$. A map belongs to $u \in W_{loc}^{k,p}(\Omega)$ 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 $\varphi \in C_c^\infty(\mathbb{R}^n)$,*

$$\int_{\mathbb{R}^n} f(x)\,\partial_i \varphi(x)\,dx = \int_{\mathbb{R}^n} s\,|x|^{-s-2} x^i \varphi(x)\,dx.$$

(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

$$\|u_i - u\|_{W^{k,p}(\Omega)} \xrightarrow{i\to\infty} 0.$$

□

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 $u \in W^{k,p}(\Omega)$, $1 \le p < \infty$. 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

$$\Omega_{-\varepsilon} := \{x \in \Omega : \mathrm{dist}\,(x, \partial\Omega) > \varepsilon\}$$

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, $\mathrm{supp}\,\eta_\varepsilon \subset B(0,\varepsilon)$.

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 $\varepsilon < \mathrm{dist}\,(\Omega', \partial\Omega)$, for any $1 \le p < \infty$[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 $|\gamma| \le k$. We conclude that

$$\|u_\varepsilon - u\|_{W^{k,p}(\Omega')} \xrightarrow{\varepsilon \to 0} 0.$$

□

Even though Proposition 13.9 is only about local approximation, it is very useful to prove properties of Sobolev function.

Lemma 13.10. *For $1 \le p < \infty$[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*

$$\partial_\alpha(f(v)) = f'(v)\, \partial_\alpha v. \tag{13.2}$$

[30] but not for $p = \infty$!

[31] we can later conclude, using Theorem 13.24, that this also holds for $p = \infty$, since then Sobolev maps are simply Lipschitz maps

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,

$$\|f(v_\varepsilon) - f(v)\|_{L^p(\Omega')} \le \|f'\|_{L^\infty(\Omega')}\ \|v_\varepsilon - v\|_{L^p(\Omega')} \xrightarrow{\varepsilon\to 0} 0.$$

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:

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

$$D|u| = \frac{u}{|u|}Du.$$

[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

[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

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 $\{u = 0\}$). □

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

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, 2014a, 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 $Q_j^\ast \subset Q_j^{\ast\ast}$ are cubes with the same center as Q_j but slightly increased sidelength (e.g. $\text{sidelength}(Q_j^\ast) = \frac{17}{16}\text{sidelength}(Q_j)$ and $\text{sidelength}(Q_j^{\ast\ast}) = \frac{18}{16}\text{sidelength}(Q_j)$) then each $Q_j^\ast, Q_j^{\ast\ast}$ is still contained in Ω. Moreover any $Q_j^\ast$ it intersects with at most $12^n - 4^n$ other cubes $Q_k^\ast$, and likewise any $Q_j^{\ast\ast}$ it intersects with at most $12^n - 4^n$ other cubes $Q_k^{\ast\ast}$

Now we can find a decomposition of unity $\eta_j \in C_c^\infty(Q_j^\ast)$ 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 $\operatorname{supp} \eta_j f \subset Q_j^\ast \subset\subset \Omega$. As in Proposition 13.9 we can then find $g_j \in C_c^\infty(Q_j^{\ast\ast})$ such that

$$\|g_j - f\|_{W^{k,p}(\Omega)} = \|g_j - f\|_{W^{k,p}(Q_j^\ast)} < 2^{-j}\varepsilon.$$

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 $\ell = 0,\ldots,k$

$$\begin{aligned}\|\nabla^\ell(u - u\varphi_R)\|_{L^p(\mathbb{R}^n)} =&\|\nabla^\ell(1-\varphi_R)u\|_{L^p(\mathbb{R}^n)}\\ \leq& C(\ell)\sum_{i=0}^{\ell}\|\nabla^i(1-\varphi_R)\nabla^{\ell-i}u\|_{L^p(\mathbb{R}^n)}\\ \leq& C(\ell)\,\|(1-\varphi_R)\nabla^\ell u\|_{L^p(\mathbb{R}^n)} + C(\ell)\sum_{i=0}^{\ell}\|\nabla^i(1-\varphi_R)\|_\infty\|\nabla^{\ell-i}u\|_{L^p(\mathbb{R}^n)}\\ \leq& C(\ell,\eta)\,\|(1-\varphi_R)\nabla^\ell u\|_{L^p(\mathbb{R}^n\setminus B(0,R))} + C(\ell,\eta)\sum_{i=0}^{\ell}R^{-i}\|u\|_{W^{k,p}(\mathbb{R}^n)}\\ \xrightarrow{R\to\infty}& 0\end{aligned}$$

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

$$\|u - u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)} \leq \frac{1}{2\ell}.$$

Next, we can choose ε_ℓ small enough such that

$$\|\eta_{\varepsilon_\ell} * (u\varphi_{R_\ell}) - u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)} \leq \frac{1}{2\ell}.$$

Thus, by triangular inequality,

$$\|u_\ell - u\|_{W^{k,p}(\mathbb{R}^n)} \leq \underbrace{\|u_\ell - u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)}}_{=\|\eta_{\varepsilon_\ell}*(u\varphi_{R_\ell})-u\varphi_{R_\ell}\|_{W^{k,p}(\mathbb{R}^n)}} + \|u\varphi_{R_\ell} - u\|_{W^{k,p}(\mathbb{R}^n)} \leq 2\frac{1}{2\ell} = \frac{1}{\ell}.$$

This proves (13.4), and thus (2) is established. □

Exercise 13.19. *Let*

$$f(x) := \begin{cases} 1 & x_1 \geq 0 \\ 0 & x_1 < 0 \end{cases}$$

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,

$$\|\Delta_h^{e_\ell} u\|_{L^\infty} \le \|\partial_\ell u\|_{L^\infty}.$$

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. $\square$

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*

$$\|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')} \le \|\partial_\ell u\|_{W^{k-1,p}(\Omega)}.$$

Moreover we have

$$\|\Delta_h^{e_\ell} u - \partial_\ell u\|_{W^{k-1,p}(\Omega')} \xrightarrow{h\to 0} 0.$$

(2) Let $u\in W^{k-1,p}(\Omega)$, $1<p\le\infty$. Assume that for any $\Omega'\subset\subset\Omega$ and any $\ell=1,\ldots,n$ there exists a constant $C(\Omega')$ such that

$$\sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')} \le C(\Omega',\ell)$$

Then we $u\in W^{k,p}_{loc}(\Omega)$, and for any $\Omega'\subset\Omega$ we have

$$\|\partial_\ell u\|_{W^{k-1,p}(\Omega')} \le \sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')}. \tag{13.6}$$

If $p=\infty$ we even have $u\in W^{k,\infty}(\Omega)$ with the estimate

$$\|\partial_\ell u\|_{W^{k-1,\infty}(\Omega)} \le \sup_{\Omega'\subset\subset\Omega}\ \sup_{|h|<\operatorname{dist}(\Omega',\partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,\infty}(\Omega')}. \tag{13.7}$$

---

[34]This can be seen by Jensens inequalty: For any $p\in[1,\infty)$,

$$(⨍_A |f|)^p \le ⨍_A |f|^p,$$

this can also be shown by Hölder's inequality. Then Fubini gets to the claim.

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 $|h| < \operatorname{dist}(\Omega', \partial\Omega)$.

$$\|\Delta_h^{e_\ell} \partial^\gamma u\|_{L^p(\Omega')} \leq \left(\int_0^1 \|u\|^p_{W^{k,p}(\Omega)}\, dt\right)^{\frac{1}{p}} = \|u\|_{W^{k,p}(\Omega)}.$$

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

$$\|\partial_\ell \partial^\gamma u - \Gamma_\varepsilon\|_{L^p(\Omega'')} < \varepsilon,$$

[35] This can be seen by Jensens inequalty: For any $p \in [1,\infty)$,

$$(⨍_A |f|)^p \leq ⨍_A |f|^p,$$

this can also be shown by Hölder's inequality. Then Fubini gets to the claim.

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

$$\sup_{|h| < \text{dist}\,(\Omega', \partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,p}(\Omega')} < \infty.$$

$$\Delta_{h_i}^{e_\ell} \partial^\gamma u \xrightarrow{i\to\infty} f_{\ell,\gamma} \quad \text{weakly in } L^p(\Omega').$$

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

$$\sup_{|h| < \text{dist}\,(\Omega', \partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,\infty}(\Omega')} \leq C(\Omega', \ell) \tag{13.9}$$

implies (by Hölders inequality) also

$$\sup_{|h| < \text{dist}\,(\Omega', \partial\Omega)} \|\Delta_h^{e_\ell} u\|_{W^{k-1,2}(\Omega')} \leq C(\Omega', \ell)$$

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 $\partial_\ell u \in W^{k-1,2}_{\rm loc}(\Omega)$.

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 $r < \frac{1}{4} \operatorname{dist}(x, \partial\Omega)$ 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,

$$\sup_{\Omega' \subset\subset \Omega}\ \sup_{|h| < \operatorname{dist}(\Omega', \partial\Omega)} \|\Delta_h^{e_\ell} u\|_{L^\infty(\Omega')} \leq [u]_{\mathrm{Lip}\,(\Omega)}.$$

From Proposition 13.23(2) we then obtain $u \in W^{1,\infty}(\Omega)$ with the claimed estimate.

For (2):

First we assume that $\Omega = B(0,1)$ is a ball, $u \in W^{1,\infty}(B(0,1))$. 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 $\Omega \subset\subset \mathbb{R}^n$ and $\partial\Omega \in C^{0,1}$. Moreover we assume Ω is path-connected.

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

$$[u]_{\mathrm{Lip}\,(B(x_i,\delta))} \leq \|Du\|_{L^\infty(B(x_i,\delta))} \leq \|Du\|_{L^\infty(\Omega)}.$$

Now, by triangular inequality

$$\begin{aligned}|u(x)-u(y)| \leq& |u(x)-u(x_0)| + |u(y)-u(x_L)| + \sum_{i=1}^{L-1} |u(x_i)-u(x_{i+1})| \\ \leq& \|Du\|_{L^\infty(\Omega)}\,(L+1)2\delta \leq C\mathcal{L}(\gamma) \leq C(\Omega)\|Du\|_{L^\infty(\Omega)}\,|x-y|.\end{aligned}$$

This implies that u is Lipschitz continuous with

$$[u]_{\mathrm{Lip}} \leq C(\Omega)\|Du\|_{L^\infty(\Omega)}.$$

which establishes the theorem. □

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 $|\gamma| \leq k$ 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

$$\|f\|_{W^{k,p}(\Omega)} \leq \limsup_i \|f_i\|_{W^{k,p}(\Omega)}.$$

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.

$$\sup_{\|u\|_{W^{k,p}(\Omega)}\leq 1} \|Eu\|_{W^{k,p}(\mathbb{R}^n)} < \infty.$$

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 $k \in \mathbb{N}$, $1 \leq p < \infty$.*

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

$$\|Eu\|_{W^{k,p}(\tilde{\Omega})} \leq C(\Omega, \tilde{\Omega}, n, k)\, \|u\|_{W^{k,p}(\Omega)}.$$

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.

From $\mathbb{R}^n_+$ to $\mathbb{R}^n$:

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,

$$E_0 u(y', y_n) := \begin{cases} u(y', y_n) & \text{if } y_n > 0 \\ \sum_{i=1}^k \sigma_i\, u(y, -i y_n) & \text{if } y_n < 0. \end{cases}$$

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 $Eu \in W^{k,p}(\mathbb{R}^n)$.

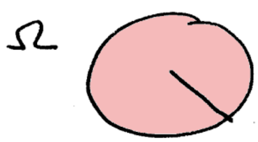
From Ω to Ω' To get $E_2 u \in W^{k,p}_0(\Omega')$ we simply take a cuttoff function $\eta \in C_c^\infty(\Omega')$, $\eta \equiv 1$ in Ω, and set
$$E_2 u := \eta E_1 u.$$
□

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),

$$\int_{\mathbb{R}^n_+} p|u(x)|^{p-2}u(x)\,\partial_n u(x)\,dx \leq C \int_{\mathbb{R}^n_+} (|u|^p + |\partial_n u|^p) \leq C\|u\|^p_{W^{1,p}(\mathbb{R}^n_+)}.$$

(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. □

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.
$$\|Tf\|_{W^{1-\frac{1}{p},p}(\partial\Omega)} \leq C(\Omega)\,\|f\|_{W^{1,p}(\Omega)}$$
- If $f \in W^{1,p}(\Omega) \cap C^0(\overline{\Omega})$ then $Tf = f\Big|_{\partial\Omega}$
- There exists a linear bounded operator $S : W^{1-\frac{1}{p},p}(\partial\Omega) \to W^{1,p}(\Omega)$,
$$\|Sf\|_{W^{1,p}(\Omega)} \leq \tilde{C}(\Omega)\,\|f\|_{W^{1-\frac{1}{p},p}(\partial\Omega)}$$
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
$$\|u\|_{L^p(\Omega)} \leq \|u\|_{W^{1,p}(\Omega)}.$$
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

$$\sup_{k\in\mathbb{N}} \|u_{k,\varepsilon}\|_{\mathrm{Lip}\,(\mathbb{R}^n)} \leq C(\varepsilon).$$

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,

$$\left|u_{k_{i_1}}(x) - u_{k_{i_1},\varepsilon_j}(x)\right| = \int_{\mathbb{R}^n} |\eta_\varepsilon(z)| \left|u_{k_{i_1}}(x-z) - u_{k_{i_1}}(x)\right| \leq \int_0^1 \int_{B(0,\varepsilon)} |\eta_\varepsilon(z)||Du_{k_{i_1}}(x-tz))|\,|z| dz\,dt$$

$$\leq \varepsilon^{1-n}\int_0^1 \left(\int_{B(0,\varepsilon)} |Du_{k_{i_1}}(x-tz))|^p\right)^{\frac{1}{p}} \varepsilon^{n-\frac{n}{p}} dz\,dt$$

Thus, by Fubini

(13.13)

$$\|u_{k_{i_1}} - u_{k_{i_1},\varepsilon_j}\|_{L^p(\mathbb{R}^n)} = \|u_{k_{i_1}} - u_{k_{i_1},\varepsilon_j}\|_{L^p(B(0,2R))} \leq \varepsilon^{1-\frac{n}{p}} \left(\int_0^1 \int_{B(0,R)} \int_{B(0,\varepsilon)} |Du_{k_{i_1}}(x-tz) dx dz dt\right)^{\frac{1}{p}}$$

$$\leq \varepsilon \|Du_{k_{i_1}}\|_{L^p(B(0,2R))}.$$

Now we claim that this leads to a Cauchy-sequence for the (non-mollified) u_{k_i}: Let $\delta > 0$.

$$\|u_{k_{i_1}} - u_{k_{i_2}}\|_{L^p(\mathbb{R}^n)} \leq \|u_{k_{i_1}} - u_{k_{i_1},\varepsilon_j}\|_{L^p(\mathbb{R}^n)} + \|u_{k_{i_1},\varepsilon_j} - u_{k_{i_2},\varepsilon_j}\|_{L^p(\mathbb{R}^n)} + \|u_{k_{i_2},\varepsilon_j} - u_{k_{i_2}}\|_{L^p(\mathbb{R}^n)}$$

$$\overset{(13.13)}{\leq} 2\,C\,\varepsilon_j \sup_k \|u_k\|_{W^{1,p}(\mathbb{R}^n)} + C(R)\|u_{k_{i_1},\varepsilon_j} - u_{k_{i_2},\varepsilon_j}\|_{L^\infty(B(2R))}$$

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)

$$\|u-(u)_I\|^p_{L^p(I)} \leq |I|^{-1}\int_I\int_I |u(x)-u(y)|^p \leq |I|^{-1}\int_0^1\int_I\int_I |u'(tx+(1-t)y)|^p\,|x-y|^p\,dx\,dydt$$

Now observe that by substituting $\tilde{y} := tx + (1-t)y$, we have

$$\begin{aligned}
&|I|^{-1}\int_0^{\frac{1}{2}}\int_I\int_I |u'(tx+(1-t)y)|^p\,|x-y|^p\,dx\,dydt\\
\leq &C(I,p)\int_0^{\frac{1}{2}}\int_I\int_I \frac{1}{1-t}|u'(\tilde{y})|^p\,d\tilde{y}\,dxdt\\
= &C(I,p)\int_0^{\frac{1}{2}}\frac{1}{1-t}\,dt\ |I|\ \|u'\|^p_{L^p(I)}\\
= &\tilde{C}(I,p)\|u'\|^p_{L^p(I)}.
\end{aligned}$$

In the same way, substituting $\tilde{x} := tx + (1-t)y$

$$\begin{aligned}
&|I|^{-1}\int_{\frac{1}{2}}^1\int_I\int_I |u'(tx+(1-t)y)|^p\,|x-y|^p\,dx\,dydt\\
\leq &\tilde{C}(I,p)\|u'\|^p_{L^p(I)}.
\end{aligned}$$

So (13.14) is established.

The Poincarè inequality says that (13.14) holds also in higher dimensions,

$$\|u-(u)_\Omega\|_{L^p(\Omega)} \le C(\Omega,p)\|\nabla u\|_{L^p(\Omega)}. \tag{13.15}$$

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, $\partial\Omega \in C^{0,1}$, $1 \le p \le \infty$.*

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 $\lambda \ge 0$.
(2) if $u \in K$ and $u \equiv const$ then $u \equiv 0$.

Then there exists a constant $C = C(K,\Omega)$ such that

$$\|u\|_{L^p(\Omega)} \le C\|\nabla u\|_{L^p(\Omega)} \quad \forall u \in K. \tag{13.16}$$

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.

$$\|u_m\|_{L^p(\Omega)} = 1, \quad \|\nabla u_m\|_{L^p(\Omega)} \le \frac{1}{m}. \tag{13.18}$$

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. □

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 $u \in W^{1,p}_0(\Omega)$

$$\|u\|_{L^p(\Omega)} \leq C(\Omega)\|\nabla u\|_{L^p(\Omega)}$$

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. □

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

$$\|u - p\|_{L^p(\Omega)} \leq C \left\|\nabla^k u\right\|_{L^p(\Omega)}$$

- *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?

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

$$|u(x_1,x_2,\ldots,x_n)|\leq\int_{\mathbb{R}}|Du(z_1,x_2,\ldots,x_n)|\,dz_1.$$

[36] The optimal constant $C(p,n)$ has actually a geometric meaning, and is related to the isoperimetric inequality, cf. [Talenti, 1976]

The same way we obtain, for any $\ell = 1, \dots, n$,

$$|u(x_1, x_2, \dots, x_n)| \leq \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell.$$

Multiplying these estimates for $\ell = 1, \dots, n$ we obtain

$$|u(x_1, x_2, \dots, x_n)|^n \leq \Pi_{\ell=1}^n \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell.$$

Now we prove the case $p = 1$, when $p^* = \frac{n}{n-1}$. We have

$$\begin{aligned}&\int_{\mathbb{R}} |u(x_1, x_2, \dots, x_n)|^{\frac{n}{n-1}} dx_1 \leq \int_{\mathbb{R}} \Pi_{\ell=1}^n \left(\int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\right)^{\frac{1}{n-1}} dx_1\\ \leq &\left(\int_{z_1\in\mathbb{R}} |Du(z_1, x_2, \dots, x_n)|\, dz_1\right)^{\frac{1}{n-1}} \int_{\mathbb{R}} \Pi_{\ell=2}^n \left(\int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\right)^{\frac{1}{n-1}} dx_1\end{aligned}$$

Now by Hölder's inequality[37]

$$\int_{\mathbb{R}} \Pi_{\ell=2}^n \left(\int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\right)^{\frac{1}{n-1}} dx_1 \leq \left(\Pi_{\ell=2}^n \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\, dx_1\right)^{\frac{1}{n-1}}$$

and thus

$$\int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 \leq \left(\int_{z_1\in\mathbb{R}} |Du(z_1, x_2, \dots, x_n)|\, dz_1\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=2}^n \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\, dx_1\right)^{\frac{1}{n-1}}$$

Now we integrate this with respect to x_2, and again by Hölder's inequality,

$$\begin{aligned}&\int_{\mathbb{R}} \int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 dx_2\\ \leq &\int_{x_2\in\mathbb{R}} \left(\int_{z_1\in\mathbb{R}} |Du(z_1, x_2, \dots, x_n)|\, dz_1\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=2}^n \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\, dx_1\right)^{\frac{1}{n-1}} dx_2\\ \leq &\left(\int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, z_2, x_3, \dots x_n)|\, dz_2\, dx_1\right)^{\frac{1}{n-1}}\\ &\quad \cdot \int_{x_2\in\mathbb{R}} \left(\int_{z_1\in\mathbb{R}} |Du(z_1, x_2, \dots, x_n)|\, dz_1\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=3}^n \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\, dx_1\right)^{\frac{1}{n-1}} dx_2\\ \leq &\left(\int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, z_2, x_3, \dots x_n)|\, dz_2\, dx_1\right)^{\frac{1}{n-1}}\\ &\quad \cdot \left(\int_{x_2\in\mathbb{R}} \int_{z_1\in\mathbb{R}} |Du(z_1, x_2, \dots, x_n)|\, dz_1\, dx_2\right)^{\frac{1}{n-1}} \left(\Pi_{\ell=3}^n \int_{x_2\in\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} |Du(x_1, x_2, \dots, z_\ell, \dots x_n)|\, dz_\ell\, dx_1, dx_2\right)\end{aligned}$$

[37] the generalized version for $k := n-1$ and all $p_i := n-1$: whenever $p_1, \dots, p_k \in [1,\infty]$ and $\sum_i \frac{1}{p_i} = 1$,

$$\int_{\mathbb{R}^d} \Pi_{i=1}^k |f_i| \leq \Pi_{i=1}^k \left(\int_{\mathbb{R}^d} |f_i|^{p_i}\right)^{\frac{1}{p_i}}$$

If $n = 2$ we are done (the $\Pi_{\ell=3}^n$-term is one). If $n \geq 3$ we see a pattern, continuing to integrate in $x_3, \ldots, x_n$ we obtain

$$\int_{\mathbb{R}} \int_{\mathbb{R}} |u(x)|^{\frac{n}{n-1}} dx_1 dx_2 \leq \Pi_{\ell=1}^n \left(\int_{\mathbb{R}^n} |Du(x_1, \ldots, x_{\ell-1}, z_\ell, x_{\ell+1} \ldots x_n)| dx_1, \ldots, x_{\ell-1}, z_\ell, x_{\ell+1} \ldots x_n) \right)^{\frac{1}{n-1}}$$
$$= \left(\int_{\mathbb{R}^n} |Du| \right)^{\frac{n}{n-1}}.$$

Taking the exponent $\frac{n-1}{n}$ on both sides we obtain

$$\|u\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq \|Du\|_{L^1(\mathbb{R}^n)} \tag{13.21}$$

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

$$|Dv| = |D|u|^\gamma| \leq \gamma |u|^{\gamma-1} |Du|,$$

thus (13.21) applied to v

$$\||u|^\gamma\|_{L^{\frac{n}{n-1}}(\mathbb{R}^n)} \leq C(\gamma) \||u|^{\gamma-1} |Du|\|_{L^1(\mathbb{R}^n)} \tag{13.22}$$

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, $p' = \frac{p}{p-1}$,

$$\||u|^{\gamma-1} |Du|\|_{L^1(\mathbb{R}^n)} \leq \||u|^{\gamma-1}\|_{L^{p'}(\mathbb{R}^n)} \|Du\|_{L^p(\mathbb{R}^n)} = \|u\|^{\gamma-1}_{L^{p'(\gamma-1)}(\mathbb{R}^n)} \|Du\|_{L^p(\mathbb{R}^n)}$$

So (13.22) becomes

$$\|u\|^\gamma_{L^{\gamma\frac{n}{n-1}}(\mathbb{R}^n)} \|u\|^{1-\gamma}_{L^{p'(\gamma-1)}(\mathbb{R}^n)} \leq C(\gamma) \|Du\|_{L^p(\mathbb{R}^n)}$$

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 $\lambda > 0$,

$$\lambda^{-\frac{n}{q}} \|u\|_{L^q(\mathbb{R}^n)} \le \lambda^{1-\frac{n}{p}} \|\nabla u\|_{L^p(\mathbb{R}^n)}.$$

Equivalently, setting $\Lambda := \|\nabla u\|_{L^q(\mathbb{R}^n)} / \|u\|_{L^q(\mathbb{R}^n)} > 0$ we obtain

$$\lambda^{0-\frac{n}{q}-(1-\frac{n}{p})} \le \Lambda \quad \forall \lambda > 0$$

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

$$\lambda^\sigma \le \Lambda \quad \forall \lambda > 0$$

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 $f \in W^{1,p}(\mathbb{R}^n)$, $1 \le p < n$. For any $q \in [p, p^*]$ we have $f \in L^q(\mathbb{R}^n)$ with the estimate*

$$\|f\|_{L^q(\mathbb{R}^n)} \le C(q,n) \left(\|f\|_{L^p(\mathbb{R}^n)} + \|Df\|_{L^p(\mathbb{R}^n)} \right).$$

Proof. We first claim that

$$\|f\|^q_{L^q(\mathbb{R}^n)} \le C(q,n,p) \left(\|f\|^p_{L^p(\mathbb{R}^N)} + \|Df\|^{p^*}_{L^p(\mathbb{R}^n)} \right). \tag{13.24}$$

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*).

$$\int_{\mathbb{R}^n} |f|^q = \int_{\mathbb{R}^n} |f|^q \chi_{|f|>1} + \int_{\mathbb{R}^n} |f|^q \chi_{|f|\le 1} \le \int_{\mathbb{R}^n} |f|^{p^*} + \int_{\mathbb{R}^n} |f|^p.$$

That is,

$$\|f\|^q_{L^q(\mathbb{R}^n)} \le \|f\|^p_{L^p(\mathbb{R}^n)} + \|f\|^{p^*}_{L^{p^*}(\mathbb{R}^n)} \precsim \|f\|^p_{L^p(\mathbb{R}^n)} + \|Df\|^{p^*}_{L^p(\mathbb{R}^n)}.$$

This proves (13.24).

How do we get the main claim? Well from (13.24) we find

$$\sup_{f \in W^{1,p}(\mathbb{R}^n):\ \|f\|_{W^{1,p}(\mathbb{R}^n)} \le 1} \|f\|_{L^q(\mathbb{R}^n)} \le \tilde{C}(p,n,q).$$

So in particular we have for any $f \in W^{1,p}(\mathbb{R}^n)$.

$$\|f / \|f\|_{W^{1,p}(\mathbb{R}^n)}\|_{L^q(\mathbb{R}^n)} \le \tilde{C}(p,n,q)$$

Thus,

$$\|f\|_{L^q(\mathbb{R}^n)} \le \tilde{C}(p,n,q)\, \|f\|_{W^{1,p}(\mathbb{R}^n)}.$$

□

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 $1 \le p < n$ we have for any $u \in W^{1,p}(\Omega)$,*

$$\|u\|_{L^{p^*}(\Omega)} \le C(\Omega) \left(\|u\|_{L^p(\Omega)} + \|Du\|_{L^p(\Omega)}\right)$$

Also, for any $q \in [p, p^*]$

$$\|u\|_{L^q(\Omega)} \le C(\Omega, q)\, \|u\|_{W^{1,p}(\Omega)}.$$

If moreover $\Omega \subset\subset \mathbb{R}^n$ and $u \in W^{1,p}_0(\Omega)$ then

$$\|u\|_{L^{p^*}(\Omega)} \le C(\Omega)\, \|Du\|_{L^p(\Omega)}.$$

Lastly, if $1 \le p < \infty$ 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 $p \ge n$)

$$\|u\|_{L^q(\Omega)} \le C(\Omega, q, p, n)\ \|u\|_{W^{1,p}(\Omega)}.$$

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

$$\|u\|_{L^{p^*}(\Omega)} \le \|Eu\|_{L^{p^*}(\mathbb{R}^n)} \lesssim \|DEu\|_{L^p(\mathbb{R}^n)} \lesssim C(\Omega)\|u\|_{W^{1,p}(\Omega)} \le C(\Omega) \left(\|u\|_{L^p(\Omega)} + \|Du\|_{L^p(\Omega)}\right).$$

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

$$\|u\|_{L^q(\Omega)} \le C(\Omega, q, \Lambda) \quad \forall u: \ \|u\|_{W^{1,p}(\Omega)} \le \Lambda.$$

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 $u \in W^{1,p}_0(\Omega)$, $\Omega \subset\subset \mathbb{R}^n$ we have

$$\|u\|_{L^p(\Omega)} \le C(\Omega)\|Du\|_{L^p(\Omega)}.$$

The last claim follows by additionall using Hölder's inequality: if $p < n$, $q \in [1, p^*]$,

$$\|u\|_{L^q(\Omega)} \le C(|\Omega|, q, p)\|u\|_{L^{p^*}(\Omega)} \le C(|\Omega|, q, p)\|u\|_{W^{1,p}(\Omega)}.$$

If $p \ge n$ and $n \ge 2$, and $q \in [p, \infty)$ we can find $1 \le r \le p$ such that $\infty > r^* = \frac{nr}{n-r} > q$. Thus, from first Hölder's inequality, then Sobolev inequality, and then again Hölder's inequality,

$$\|u\|_{L^q(\Omega)} \le C(|\Omega|, q, p)\|u\|_{L^{r^*}(\Omega)} \le C(|\Omega|, q, p)\|u\|_{W^{1,r}(\Omega)} \le C(|\Omega|, q, p)\|u\|_{W^{1,p}(\Omega)}.$$

So the only remaining case is $n = 1$. Then by assumption $\Omega = (a, b)$. From the fundamental theorem of calulus

$$|f(x) - f(y)| \le \int_{(a,b)} |f'(z)| dz \quad \forall x, y \in (a, b).$$

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 $q, p \in [1, \infty]$.

$$\|f\|_{L^q((a,b))} \leq C(a,b,p,q) \|f\|_{W^{1,p}((a,b))}.$$

□

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,

$$\|u\|_{W^{\ell,q}(\Omega)} \precsim C(k,\ell,p,q) \, \|u\|_{W^{k,p}(\Omega)} \tag{13.26}$$

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 $q \le r_N$, 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 $|\gamma| \le \ell$.

So let $1 \le q < p^*$ (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 $p \ge n$ 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))

$$\sup_j \|u_{i_j}\|_{L^q(E)} \le |E|^{\frac{1}{q}-\frac{1}{r}} \sup_j \|u_{i_j}\|_{L^r(E)} \le \delta^{\frac{1}{q}-\frac{1}{r}}\Lambda.$$

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 subseqeunce 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. □

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

$$\|D^j u\|_{L^p(\mathbb{R}^n)} \leq C \, \|D^m u\|^\alpha_{L^r(\mathbb{R}^n)} \, \|u\|^{1-\alpha}_{L^q(\mathbb{R}^n)}.$$

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

$$\|D^j u\|_{L^p(\mathbb{R}^n)} \leq C \left(\|D^m u\|_{L^r(\mathbb{R}^n)} + \|u\|_{L^q(\mathbb{R}^n)}\right)$$

Indeed, then we can obtain the claim by scaling: Apply the inequality to $u(\lambda\cdot)$, then we get

$$\lambda^{-\frac{n}{p}} \lambda^j \|D^j u\|_{L^p(\mathbb{R}^n)} \leq C \left(\lambda^{-\frac{n}{r}} \lambda^m \|D^m u\|_{L^r(\mathbb{R}^n)} + \lambda^{-\frac{n}{q}} \|u\|_{L^q(\mathbb{R}^n)}\right)$$

That is

$$\|D^j u\|_{L^p(\mathbb{R}^n)} \leq C \left(\lambda^{-\frac{n}{r}+\frac{n}{p}+m-j} \|D^m u\|_{L^r(\mathbb{R}^n)} + \lambda^{\frac{n}{p}-j-\frac{n}{q}} \|u\|_{L^q(\mathbb{R}^n)}\right)$$

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!

So compute

$$\min_{\lambda \geq 0} \left(\lambda^{-\frac{n}{r}+\frac{n}{p}+m-j} \|D^m u\|_{L^r(\mathbb{R}^n)} + \lambda^{\frac{n}{p}-j-\frac{n}{q}} \|u\|_{L^q(\mathbb{R}^n)} \right).$$

(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. $\square$

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

$$\int_{B^n} \det(D\psi^1|D\psi^2|\ldots|D\psi^n) = 0$$

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 $\operatorname{rank} D\Phi \leq n-1$ (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 <u>need to show</u> is that <u>$\deg(g_1) = \deg(g_0)$</u> 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

$$H(x) := \begin{cases} g_{2|x|-1}(x/|x|) & |x| \in [\frac{1}{2}, 1] \\ G_0(2x) & |x| \leq \frac{1}{2}. \end{cases}$$

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 rank $DH(z) \leq n-1$ for $\frac{1}{2} \leq |z| \leq 1$ 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 $\tilde{\Phi} : \mathbb{R}^n \to \mathbb{S}^{n-1}$

$$\tilde{\Phi}(x) := \begin{cases} \frac{x}{|x|} & |x| \geq 1 \\ \Phi(x) & |x| < 1, \end{cases}$$

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 $|x| \geq 2$ (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 $|x| \geq 2$.

Considering $\psi(x) := \psi(2x)$ have found a counterexample to Proposition 14.1. □

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.

□

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

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \leq \fint_{B(x,2\tilde{\delta})} |\Phi_\delta(x) - \Phi(y)|dy.$$

Now,

$$\Phi_\delta(x) = \int_{B(x,2\tilde{\delta})} \delta^{-n}\eta(z/\delta)\Phi(z)dz.$$

Thus (recall $\int \eta = 1$)

$$|\Phi_\delta(x) - \Phi(y)| \leq C \fint_{B(x,2\tilde{\delta})} |\Phi(z) - \Phi(y)|dz,$$

that is

$$\operatorname{dist}(\Phi_\delta(x), \mathbb{S}^{n-1}) \leq C \fint_{B(x,2\tilde{\delta})} \fint_{B(x,2\tilde{\delta})} |\Phi(z) - \Phi(y)|dz\, dy.$$

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), $|t\Phi(x)+(1-t)x| \ge |x| - t|\Phi(x) - x| \ge 1 - t\varepsilon > 0$ 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 $D\Phi \le n-1$. 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 $1 \le p < \infty$ 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 $1 \le p < \infty$. 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

$$B_\varepsilon(\mathcal{M}) := \{x \in \mathbb{R}^m : \quad \text{dist}\,(x, \mathcal{M}) < \varepsilon\}.$$

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

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) \leq |g_\delta(x) - g(y)| \quad \forall y \text{ close enough to } \Omega. \tag{14.6}$$

Integrating (14.4) with respect to y in $B(x, 2\delta)$ we find

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) \leq \fint_{B(x,2\delta} |g_\delta(x) - g(y)| dy.$$

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

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) \leq C \fint_{B(x,2\delta)} \fint_{B(x,2\delta)} |g(z) - g(y)| dz\, dy.$$

Triangular inequality then shows

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) \leq C \fint_{B(x,2\delta)} |g(z) - (g)_{B(x,2\delta)}| dz.$$

Poincaré inequality and Hölder's inequality (observe the power of δ!)

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) \leq C \delta^{1-n+n\frac{p}{p-1}n} \|\nabla g\|_{L^p(B(x,2\delta))}.$$

Since $p \geq n$ we have $1 - n + n\frac{p}{p-1}n \geq 0$ and thus

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) \leq C \|\nabla g\|_{L^p(B(x,2\delta))}.$$

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 $\delta \ll 1$,

$$\operatorname{dist}(g_\delta(x), \mathcal{M}) < \varepsilon,$$

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 $\pi_{\mathcal{M}} g \equiv g$.

That is, any $W^{1,p}$-map, $p \geq n$, is smoothly approximable, which is what we needed to show. □

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.

[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!

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]

$$u_{i_j}(t, \cdot) - u(t, \cdot) \xrightarrow{j \to \infty} 0 \quad \text{in } W^{1,p}\left(\mathbb{R}^{n-1}\right)$$

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}

$$F_{k_j}(t) \xrightarrow{j \to \infty} 0 \quad \mathcal{L}^1\text{-a.e. } t.$$

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

$$\|u_{k_j}(t, \cdot) - u(t, \cdot)\|_{W^{1,p}(\mathbb{R}^{n-1})} \xrightarrow{j \to \infty} 0.$$

[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 $\mathbb{R}^n$!

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!) $u \not\in W^{1,1}$. 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 $\phi \in C^1_c(\Omega, \mathbb{R}^n)$ the *divergence* $\operatorname{div} \phi \in C^0_c(\Omega)$ 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*

$$\Lambda := \sup\left\{\int_\Omega f \operatorname{div} \phi d\mathcal{L}^n : \quad \phi \in C^1_c(\Omega,\mathbb{R}^n), \|\phi\|_{L^{p'}(\Omega)} \leq 1\right\} < \infty.$$

In that case there exists a constant such that $C^{-1}\|Df\|_{L^p} \leq \Lambda \leq C\|Df\|_{L^p}$.

Proof. $\underline{\Rightarrow}$ Assume $f \in W^{1,p}(\Omega)$ then

$$\int f \operatorname{div} \phi = \int Df \cdot \phi \leq \|Df\|_{L^p} \|\phi\|_{L^{p'}}$$

$\underline{\Leftarrow}$ Let $\phi = (0,\ldots,0,\varphi,0,\ldots,0)$ (where φ is at the j-th position then we get that

$$\varphi \mapsto \int_\Omega f \partial_j \varphi \leq \Lambda \|\varphi\|_{L^{p'}(\mathbb{R}^n)}.$$

By the Riesz representation theorem (here we need that $p' < \infty$, so $p > 1$) we find $g \in L^p(\mathbb{R}^n)$, $\|g\|_{L^p} \leq \Lambda$ 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
$$\sup \left\{ \int_\Omega f \operatorname{div} \phi d\mathcal{L}^n : \quad \phi \in C_c^1(\Omega, \mathbb{R}^n), |\phi(x)| \leq 1 \text{ in } \Omega \right\} < \infty.$$
(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 $\Omega' \subset\subset \Omega$,
$$\sup \left\{ \int_{\Omega'} f \operatorname{div} \phi d\mathcal{L}^n : \quad \phi \in C_c^1(\Omega', \mathbb{R}^n), |\phi(x)| \leq 1 \text{ in } \Omega' \right\} < \infty.$$
(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 $\phi \in C_c^1(\Omega,\mathbb{R}^n)$

$$\int_\Omega f \operatorname{div} \phi = \int_\Omega \nabla f \cdot \phi \leq \|\phi\|_{L^\infty} \|\nabla f\|_{L^1}.$$

□

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

$$\int_E \operatorname{div}(\phi) = \int_{\partial E} \nu \cdot \phi \leq \mathcal{H}^{n-1}(\partial E)\, \|\phi\|_{L^\infty(\partial E)} \leq \mathcal{H}^{n-1}(\partial E) \|\phi\|_{L^\infty(\mathbb{R}^n)} \quad \forall \phi \in C_c^1(\mathbb{R}^n).$$

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

$$\mathrm{Per}(E) = \mathcal{H}^{n-1}(\partial E).$$

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 $f_k \to f \in L^1_{loc}(\Omega)$.*

$$\|Df\|(\Omega) \leq \liminf_{k\to\infty} \|Df_k\|(\Omega),$$

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 $\phi \in C^1_c(\Omega, \mathbb{R}^n)$. 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. □

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

$$f_{k_j} \xrightarrow{j\to\infty} f \quad \text{in } L^1(U).$$

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 subseqeunce) g_{k_j} which converges strongly in $L^1(\mathcal{U})$, and so f_k converges in $L^1(U)$. □

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 $u_k \in C^\infty_c(\mathbb{R}^n)$ 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

$$\int_E \mathrm{div}\, \phi = \int_{E\cap B(x,r)} \mathrm{div}\, \phi.$$

Either, if $B(x,r) \subset E$

$$\int_E \mathrm{div}\, \phi = \int_{B(x,r)} \mathrm{div}\, \phi = 0,$$

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.*

- *$\mathcal{H}^{n-1}(B) \leq C\|\partial E\|(B)$ 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*
$$\int_E \operatorname{div}\phi = \int_{\partial_* E} \phi \cdot \nu_E d\mathcal{H}^{n-1} \quad \forall \phi \in C_c^1(\mathbb{R}^n, \mathbb{R}^n).$$

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)

Email address: armin@pitt.edu